

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2006, 04:27:56 ; Search time 504 Seconds
(without alignments)
9801.046 Million cell updates/sec

Title: US-10-783-519-1
Perfect score: 2640
Sequence: 1 csgcgtcgccgagcggcag.....gcaaaaaaaaaaaaaaaaaa 2640

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /EMC_Celerra_SID33/ptodata/2/ina/1 COMB.seq.*
2: /EMC_Celerra_SID33/ptodata/2/ina/5 COMB.seq.*
3: /EMC_Celerra_SID33/ptodata/2/ina/6A COMB.seq.*
4: /EMC_Celerra_SID33/ptodata/2/ina/6B COMB.seq.*
5: /EMC_Celerra_SID33/ptodata/2/ina/7 COMB.seq.*
6: /EMC_Celerra_SID33/ptodata/2/ina/H COMB.seq.*
7: /EMC_Celerra_SID33/ptodata/2/ina/PCTUS COMB.seq.*
8: /EMC_Celerra_SID33/ptodata/2/ina/PP COMB.seq.*
9: /EMC_Celerra_SID33/ptodata/2/ina/RE COMB.seq.*
10: /EMC_Celerra_SID33/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|---------------------|
| 1 | 2640 | 100.0 | 2640 | 3 | US-08-684-932A-37 |
| 2 | 2640 | 100.0 | 2640 | 3 | US-09-618-304B-1 |
| 3 | 2640 | 100.0 | 2655 | 3 | US-09-016-434-1094 |
| 4 | 2640 | 100.0 | 2655 | 4 | US-09-023-655-916 |
| 5 | 2640 | 100.0 | 2655 | 4 | US-09-880-107-3363 |
| 6 | 2580 | 97.7 | 2654 | 3 | US-09-949-016-2404 |
| 7 | 2580 | 97.7 | 2654 | 3 | US-09-949-016-2405 |
| 8 | 2553.4 | 96.7 | 2706 | 3 | US-09-949-016-2975 |
| 9 | 2553.4 | 96.7 | 2706 | 3 | US-09-949-016-2976 |
| 10 | 2051 | 77.7 | 2457 | 3 | US-09-949-016-1330 |
| 11 | 2051 | 77.7 | 2457 | 3 | US-09-949-016-1331 |
| 12 | 741.2 | 28.1 | 2945 | 10 | 5196526-2 |
| 13 | 740.8 | 28.1 | 2568 | 10 | 5196526-3 |
| 14 | 618.6 | 23.4 | 3062 | 3 | US-09-799-451-895 |
| 15 | 554.6 | 21.0 | 4132 | 3 | US-09-566-921-83 |
| 16 | 553 | 20.9 | 3033 | 3 | US-09-016-434-1162 |
| 17 | 502.2 | 19.0 | 2339 | 3 | US-09-495-050A-26 |
| 18 | 486.6 | 18.4 | 493 | 4 | US-09-880-107-1291 |
| 19 | 301.6 | 11.4 | 1724 | 3 | US-09-270-767-11642 |
| 20 | 247.8 | 9.4 | 11917 | 3 | US-09-949-016-13072 |
| 21 | 247.8 | 9.4 | 11917 | 3 | US-09-949-016-13073 |
| 22 | 247.8 | 9.4 | 15061 | 3 | US-09-949-016-14717 |
| 23 | 247.8 | 9.4 | 15061 | 3 | US-09-949-016-14718 |

24 247.8 9.4 15859 3 US-09-949-016-14146 Sequence 14146, A
25 247.8 9.4 15859 3 US-09-949-016-14147 Sequence 14147, A
c 26 239.2 9.1 1145 3 US-09-270-767-12941 Sequence 12941, A
c 27 223.2 8.5 847 3 US-09-270-767-28814 Sequence 28814, A
28 221.8 8.4 601 3 US-09-949-016-83496 Sequence 83496, A
29 221.8 8.4 601 3 US-09-949-016-83505 Sequence 83505, A
30 221.8 8.4 601 3 US-09-949-016-107576 Sequence 107576, A
31 221.8 8.4 601 3 US-09-949-016-107585 Sequence 107585, A
32 214.2 8.1 2037 4 US-10-094-749-572 Sequence 572, App
33 214.2 8.1 601 3 US-09-949-016-46215 Sequence 46215, A
34 214.2 8.1 601 3 US-09-949-016-46222 Sequence 46222, A
35 214.2 8.1 601 3 US-09-949-016-83501 Sequence 83501, A
36 214.2 8.1 601 3 US-09-949-016-83510 Sequence 83510, A
37 214.2 8.1 601 3 US-09-949-016-107581 Sequence 107581, A
38 214.2 8.1 601 3 US-09-949-016-107590 Sequence 107590, A
c 39 166.6 6.3 762 3 US-09-270-767-11453 Sequence 11453, A
40 164.8 6.2 1351 3 US-09-270-767-1513 Sequence 1513, App
41 164.8 6.2 1351 3 US-09-270-767-16795 Sequence 16795, A
42 164.4 6.2 601 3 US-09-949-016-46216 Sequence 46216, A
43 164.4 6.2 601 3 US-09-949-016-46223 Sequence 46223, A
44 164.4 6.2 601 3 US-09-949-016-83502 Sequence 83502, A
45 164.4 6.2 601 3 US-09-949-016-83511 Sequence 83511, A

ALIGNMENTS

RESULT 1
US-08-684-932A-37
Sequence 37, Application US/08684932A
Patent No. 6403304
GENERAL INFORMATION:
APPLICANT: Stashenko, Philip
APPLICANT: Li, Yi-Ping
APPLICANT: Wucherpfermig, Anne L.
TITLE OF INVENTION: HUMAN OSTEOCLAST-SPECIFIC AND -RELATED
TITLE OF INVENTION: DNA SEQUENCES
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/684,932A
FILING DATE: 19-JUL-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: PDC92-02FM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 2640 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 58..2523
US-08-684-932A-37

| Query Match | | | | | | | | | |
|---|------|----------------------|--|--------------------|----------------------------------|----|--|--|--|
| Best Local Similarity 100.0%; Score 2640; DB 3; Length 2640; | | | | | | | | | |
| Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | | | |
| QY | 1 | CGGGTGC | CGGACGGG | CAGCCAGC | GAGGCGGCGGCGCAGCACACCCGGGGACCATG | 60 | | | |
| DB | 1 | CGGGTGC | CGGACGGG | CAGCCAGC | GAGGCGGCGGCGCAGCACACCCGGGGACCATG | 60 | | | |
| QY | 61 | GGCTCCATGTTT | CGGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTTGCCCA | CAGCGGCT | 120 | | | | |
| DB | 61 | GGCTCCATGTTT | CGGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTTGCCCA | CAGCGGCT | 120 | | | | |
| QY | 121 | GCCTACACCTG | CGTGAAGTGGGCGAGCTGGCCCTGGTGAAGTTT | CAAGACCTCAAC | 180 | | | | |
| DB | 121 | GCCTACACCTG | CGTGAAGTGGGCGAGCTGGCCCTGGTGAAGTTT | CAAGACCTCAAC | 180 | | | | |
| QY | 181 | GCCTCGGTGAGCGGCTT | CCAGACGCTTTGTGGTTGATGTTTGGCGCTGTGAGGAGCTG | 240 | | | | | |
| DB | 181 | GCCTCGGTGAGCGGCTT | CCAGACGCTTTGTGGTTGATGTTTGGCGCTGTGAGGAGCTG | 240 | | | | | |
| QY | 241 | GAGAGACCTT | CACCTTCTTGACAGGAGGTGGCGGCTGGGCTGGTCTGCCCCCG | 300 | | | | | |
| DB | 241 | GAGAGACCTT | CACCTTCTTGACAGGAGGTGGCGGCTGGGCTGGTCTGCCCCCG | 300 | | | | | |
| QY | 301 | CCAAAGGGGAGGCTG | CGGCACCCCAACCCCGGACCTGCTGGGCATCCACAGGAGGAGCG | 360 | | | | | |
| DB | 301 | CCAAAGGGGAGGCTG | CGGCACCCCAACCCCGGACCTGCTGGGCATCCACAGGAGGAGCG | 360 | | | | | |
| QY | 361 | GAGGCTTGCCCAAGGAGCTG | CGGAGTGTGGGGGMAACAGACAGGCGCTTGCGGGGCCAG | 420 | | | | | |
| DB | 361 | GAGGCTTGCCCAAGGAGCTG | CGGAGTGTGGGGGMAACAGACAGGCGCTTGCGGGGCCAG | 420 | | | | | |
| QY | 421 | CTGCACAGCTG | CAGCTCCACGCGCGCTGCTAGCCAGGCGCATGAACCTCAGCTGGCA | 480 | | | | | |
| DB | 421 | CTGCACAGCTG | CAGCTCCACGCGCGCTGCTAGCCAGGCGCATGAACCTCAGCTGGCA | 480 | | | | | |
| QY | 481 | GCCGCCACACAGATG | GGGCTCAGAGAGACGCCCCCTGCTCAGGCGCCCCGGGGGGCCG | 540 | | | | | |
| DB | 481 | GCCGCCACACAGATG | GGGCTCAGAGAGACGCCCCCTGCTCAGGCGCCCCGGGGGGCCG | 540 | | | | | |
| QY | 541 | CACAGGACCTGAGGGTCAA | CTTTGTGGCAGGTCGCTGGAGCCCCCAAGGCCCTTGCC | 600 | | | | | |
| DB | 541 | CACAGGACCTGAGGGTCAA | CTTTGTGGCAGGTCGCTGGAGCCCCCAAGGCCCTTGCC | 600 | | | | | |
| QY | 601 | CTAGAGCGCTGCT | CTGAGGGCTGCGCGGCTCTCATTCGCGAGCTTCAGGAGCTG | 660 | | | | | |
| DB | 601 | CTAGAGCGCTGCT | CTGAGGGCTGCGCGGCTCTCATTCGCGAGCTTCAGGAGCTG | 660 | | | | | |
| QY | 661 | GAGCAGCGCTGGAGCAC | CCCCGTGACGGGCGAGCCAGCCACGTTGGATGACCTTCTCATC | 720 | | | | | |
| DB | 661 | GAGCAGCGCTGGAGCAC | CCCCGTGACGGGCGAGCCAGCCACGTTGGATGACCTTCTCATC | 720 | | | | | |
| QY | 721 | TCCTACTGGGGTGAGCAGAT | TCGACAGAGATCCGCAAGATCA | CGGACTGCTTCCATGCTC | 780 | | | | |
| DB | 721 | TCCTACTGGGGTGAGCAGAT | TCGACAGAGATCCGCAAGATCA | CGGACTGCTTCCATGCTC | 780 | | | | |
| QY | 781 | CAGCTCTCCCTTTCTG | CAGAGGAGGCGCGCTCGGGCCCTGCGAGCTGCGAA | 840 | | | | | |
| DB | 781 | CAGCTCTCCCTTTCTG | CAGAGGAGGCGCGCTCGGGCCCTGCGAGCTGCGAA | 840 | | | | | |
| QY | 841 | CAGCAGGCGCAGGAGCTG | CAGGAGTCTCGGGGAGACAGCGGTTCTTGAGCCAGGTTG | 900 | | | | | |
| DB | 841 | CAGCAGGCGCAGGAGCTG | CAGGAGTCTCGGGGAGACAGCGGTTCTTGAGCCAGGTTG | 900 | | | | | |
| QY | 901 | CTAGGCGGGTGCTGACGCTG | CTGCGCCAGGGGAGGTGAGGTCCACAGATGAAGGCC | 960 | | | | | |
| DB | 901 | CTAGGCGGGTGCTGACGCTG | CTGCGCCAGGGGAGGTGAGGTCCACAGATGAAGGCC | 960 | | | | | |
| QY | 961 | GTCTACTGCGCTGAA | CCAGTGCAGGTGAGCACCAAGTGCCTCATTTGCCGAG | 1020 | | | | | |
| DB | 961 | GTCTACTGCGCTGAA | CCAGTGCAGGTGAGCACCAAGTGCCTCATTTGCCGAG | 1020 | | | | | |
| QY | 1021 | GCCTGGTGTCTGTG | CGAGACCTTGCCCGCTCGACGAGGCGCTTGCGGACAGCTCGATG | 1080 | | | | | |


```
Db 2101 GACGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGGACGGGGCCCTGGATGATGAA 2160
Qy 2161 GAGGAGCCGAGCTCTCTCCCTCCGAGGTGCTCATGCAACAGGCCATCCACACCATCGAG 2220
Db 2161 GAGGAGCCGAGCTCTCTCCCTCCGAGGTGCTCATGCAACAGGCCATCCACACCATCGAG 2220
Qy 2221 TTCTGCTGGCTGCTCTCCACACCGCTCTCTACCTGGCTGCTGGCCCTGAGGCTG 2280
Db 2221 TTCTGCTGGCTGCTCTCCACACCGCTCTCTACCTGGCTGCTGGCCCTGAGGCTG 2280
Qy 2281 GCCACGCCAGCTGCTCCAGAGTTCTGTGGGCCATGCTGATGCGCATAGGCCCTGGGCCCTG 2340
Db 2281 GCCACGCCAGCTGCTCCAGAGTTCTGTGGGCCATGCTGATGCGCATAGGCCCTGGGCCCTG 2340
Qy 2341 GCGCGGAGGTGGCTGGCGCTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
Db 2341 GCGCGGAGGTGGCTGGCGCTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
Qy 2401 ATGACCGTGTATCTGCTGTGTGATGAGGAGTCTCTAGCCCTTCTGCAAGCTGAGTCCC 2460
Db 2401 ATGACCGTGTATCTGCTGTGTGATGAGGAGTCTCTAGCCCTTCTGCAAGCTGAGTCCC 2460
Qy 2461 CTGCACTGGGTGAATTCAGAACAAAGTTCTACTCAGGCAAGGCTTCTGCAAGCTGAGTCCC 2520
Db 2461 CTGCACTGGGTGAATTCAGAACAAAGTTCTACTCAGGCAAGGCTTCTGCAAGCTGAGTCCC 2520
Qy 2521 TTCACTTCTGCTGCAAGTACTAGGCCCCACTGCAAGTCTGCAAGTCTGCAAGTCTGCTTCT 2580
Db 2521 TTCACTTCTGCTGCAAGTACTAGGCCCCACTGCAAGTCTGCAAGTCTGCAAGTCTGCTTCT 2580
Qy 2581 GACCTTGAGGAGGAGGATTAAGAGCTGCGCCCTGGGCAAGGAGGAGGAGGAGGAGGAGGAG 2640
Db 2581 GACCTTGAGGAGGAGGATTAAGAGCTGCGCCCTGGGCAAGGAGGAGGAGGAGGAGGAGGAG 2640

RESULT 2
US-09-618-304B-1
; Sequence 1, Application US/09618304B
; Patent No. 6777537
; GENERAL INFORMATION:
; APPLICANT: Stashenko, Philip
; APPLICANT: Li, Yi-Ping
; TITLE OF INVENTION: Osteoclast Proton Pump Subunit
; FILE REFERENCE: 1564.1006-001
; CURRENT APPLICATION NUMBER: US/09/618,304B
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 08/605,378
; PRIOR FILING DATE: 1996-02-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2640
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-618-304B-1

Query Match 100.0%; Score 2640; DB 3; Length 2640;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGCGTGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 60
Db 1 CCGCGTGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 60
Qy 61 GGTCTCATGTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 61 GGTCTCATGTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Qy 121 GCTACACCTGCTGAGTGGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 180
Db 121 GCTACACCTGCTGAGTGGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 180
Qy 181 GCTCGGTGAGGCGCTTCAGAGACGCTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 240
```

```
Db 181 GCTCGGTGAGGCGCTTCAGAGACGCTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 240
Qy 241 GAGAAGACCTTACCTTCTCGAGGAGGAGTGGCGGGCTGGGCTGGTCTCTGCCCCCG 300
Db 241 GAGAAGACCTTACCTTCTCGAGGAGGAGTGGCGGGCTGGGCTGGTCTCTGCCCCCG 300
Qy 301 CCAAGGGAGGAGTGGCGGACCCCAACCCCGGACCTGCTGGGATCAGAGGAGAGC 360
Db 301 CCAAGGGAGGAGTGGCGGACCCCAACCCCGGACCTGCTGGGATCAGAGGAGAGC 360
Qy 361 GAGCGCTGGCCAGAGCTGGGATGTGGGGCAACAGCAGGACCTTGGGGGCCAG 420
Db 361 GAGCGCTGGCCAGAGCTGGGATGTGGGGCAACAGCAGGACCTTGGGGGCCAG 420
Qy 421 CTGCAACAGCTCAGCTCCACCGCGCTGCTACGCGAGGCGCATGAACCTCAGCTGGCA 480
Db 421 CTGCAACAGCTCAGCTCCACCGCGCTGCTACGCGAGGCGCATGAACCTCAGCTGGCA 480
Qy 481 GCGGCCACACAGATGGGCTCAGAGAGAGCGCCCTGCTCCAGGCGCCCGGGGGCCG 540
Db 481 GCGGCCACACAGATGGGCTCAGAGAGAGCGCCCTGCTCCAGGCGCCCGGGGGCCG 540
Qy 541 CACCAAGGACCTGAGGCTCAACTTTGTGGCAGGTGCGGTGGAGCCCAAGGCGCTTGGCC 600
Db 541 CACCAAGGACCTGAGGCTCAACTTTGTGGCAGGTGCGGTGGAGCCCAAGGCGCTTGGCC 600
Qy 601 CTAGAGCGCTCTCTGAGGCGCTTCCCTCAITTCAGCTTTCAGGAGCTG 660
Db 601 CTAGAGCGCTCTCTGAGGCGCTTCCCTCAITTCAGCTTTCAGGAGCTG 660
Qy 661 GAGCAGCGCTGAGGACCCCTGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 720
Db 661 GAGCAGCGCTGAGGACCCCTGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 720
Qy 721 TCCTACTGGGTGAGCAGATCGGACAGATCGGACAGATCGGACAGATCGGACAGATCGGACAG 780
Db 721 TCCTACTGGGTGAGCAGATCGGACAGATCGGACAGATCGGACAGATCGGACAGATCGGACAG 780
Qy 781 CAGCTTCTCCCTTTCTGAGGAGGAGGCGGCTCGGGGCGCTCGGAGCGCTGAGCAGCTGCAA 840
Db 781 CAGCTTCTCCCTTTCTGAGGAGGAGGCGGCTCGGGGCGCTCGGAGCGCTGAGCAGCTGCAA 840
Qy 841 CAGCAGAGCAGAGCTGAGAGGTCTCGGGGAGAGAGAGGCTTCTGAGCAGGAGT 900
Db 841 CAGCAGAGCAGAGCTGAGAGGTCTCGGGGAGAGAGAGGCTTCTGAGCAGGAGT 900
Qy 901 CTAGCGCGGTGCTGAGCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
Db 901 CTAGCGCGGTGCTGAGCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
Qy 961 GTGTACTTGGCGCTGAAACAGTGCAGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1020
Db 961 GTGTACTTGGCGCTGAAACAGTGCAGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1020
Qy 1021 GCTGCTGCTGCTGAGAGCTGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Db 1021 GCTGCTGCTGCTGAGAGCTGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Qy 1081 GAGGAGGAGTGAAGTGGCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Db 1081 GAGGAGGAGTGAAGTGGCTGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Qy 1141 ATCCGCAACCAACCGCTTACCGGCGAGCTTTCAGGAGGAGTGGTGGATCGCTTACGCGCTGG 1200
Db 1141 ATCCGCAACCAACCGCTTACCGGCGAGCTTTCAGGAGGAGTGGTGGATCGCTTACGCGCTGG 1200
Qy 1201 CGCTACAGGAGTCAACCCCGCTCCCTACACATCATCCTTCCCTTCTGTTTGT 1260
Db 1201 CGCTACAGGAGTCAACCCCGCTCCCTACACATCATCCTTCCCTTCTGTTTGT 1260
Qy 1261 GTGATGTTGGGGAGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
```

| | | | |
|----|------|--|------|
| Db | 1361 | GTGATGTTGGGGAGTGGGCCACGGGGTGTCTCATGTTCTCTTCGCGCCCTGGCCGANGTC | 1332 |
| Qy | 1321 | CTTGGAGAAACGACCGGCTGTGAAGCCGCGCAGAACAGAGATCTGGCAGACTTTCTTC | 1380 |
| Db | 1321 | CTTGGGAGAACGACCGGCTGTGAAGCCGCGCAGAACGAGATCTGGCAGACTTTCTTC | 1380 |
| Qy | 1381 | AGGGCCGCTACCTGCTCTGCTTTATGGGCTGTTCCTCATCTACACCGGCTTCATCTAC | 1440 |
| Db | 1381 | AGGGCCGCTACCTGCTCTGCTTTATGGGCTGTTCCTCATCTACACCGGCTTCATCTAC | 1440 |
| Qy | 1441 | AACGAGTGCTCAGTCGGGCCACAGACATCTTCCCTCGGGCTGGAGTGTGGCGCCATG | 1500 |
| Db | 1441 | AACGAGTGCTCAGTCGGGCCACAGACATCTTCCCTCGGGCTGGAGTGTGGCGCCATG | 1500 |
| Qy | 1501 | GCCAAACAGTCTGGCTGGAGTGATGCACTCTCTGGCCCGCAGCACAGATGCTTACCTCGAT | 1560 |
| Db | 1501 | GCCAAACAGTCTGGCTGGAGTGATGCACTCTCTGGCCCGCAGCACAGATGCTTACCTCGAT | 1560 |
| Qy | 1561 | CCCAACGTCACCGGTGTTCTCTGGGACCCCTACCCCTTTGGCATCGATCTATTTGGAGC | 1620 |
| Db | 1561 | CCCAACGTCACCGGTGTTCTCTGGGACCCCTACCCCTTTGGCATCGATCTATTTGGAGC | 1620 |
| Qy | 1621 | CTGGCTGCCAACACACTTGAGCTTCTCAAATCTCTTAAGATGAAGATGTCGTCATCTTC | 1680 |
| Db | 1621 | CTGGCTGCCAACACACTTGAGCTTCTCAAATCTCTCTCAAGATGAAGATGTCGTCATCTTC | 1680 |
| Qy | 1681 | GGCGTCTGTCACATGGGCTTTGGGCTGTCTCGAGTCTTCAACACAGTGCACTTTGGC | 1740 |
| Db | 1681 | GGCGTCTGTCACATGGGCTTTGGGCTGTCTCGAGTCTTCAACACAGTGCACTTTGGC | 1740 |
| Qy | 1741 | CAGAGGCACCGGCTGTCTGGAGACGCTGCGGAGCTCACCTTCTGTCTGGACCTTCTTC | 1800 |
| Db | 1741 | CAGAGGCACCGGCTGTCTGGAGACGCTGCGGAGCTCACCTTCTGTCTGGACCTTCTTC | 1800 |
| Qy | 1801 | GTTTACCTCTGTGTTCTTAGTCACTCAAGTGGCTGTGTGTCTGGGCTGCCAGGGCGCC | 1860 |
| Db | 1801 | GTTTACCTCTGTGTTCTTAGTCACTCAAGTGGCTGTGTGTCTGGGCTGCCAGGGCGCC | 1860 |
| Qy | 1861 | TCGCCCCAGCATCTCATCTTCAATCAATGTTCTCTTCCACAGCCGCCACGCAAC | 1920 |
| Db | 1861 | TCGCCCCAGCATCTCATCTTCAATCAATGTTCTCTTCCACAGCCGCCACGCAAC | 1920 |
| Qy | 1921 | AGGCTGCTCTACCCCGCGCAGGAGTGCTCCAGGCCACGCTGTGTGTCTCTGGGCTTGGCC | 1980 |
| Db | 1921 | AGGCTGCTCTACCCCGCGCAGGAGTGCTCCAGGCCACGCTGTGTGTCTCTGGGCTTGGCC | 1980 |
| Qy | 1981 | ATGTTGCCATCTCTGTCTTTGGACACCCCTGCACTGTGTGACCGCCACCGCGCGCC | 2040 |
| Db | 1981 | ATGTTGCCATCTCTGTCTTTGGACACCCCTGCACTGTGTGACCGCCACCGCGCGCC | 2040 |
| Qy | 2041 | CTCGGAGGAGGCCCGCTGACCGACAGGAGAAACAGCCGGTGTCTGGACCTGCGCT | 2100 |
| Db | 2041 | CTCGGAGGAGGCCCGCTGACCGACAGGAGAAACAGCCGGTGTCTGGACCTGCGCT | 2100 |
| Qy | 2101 | GACGCATCTGTGAATGGCTGGAGCTCCGATGAGAAAGGCAGGGGCGCTGGATGATGAA | 2160 |
| Db | 2101 | GACGCATCTGTGAATGGCTGGAGCTCCGATGAGAAAGGCAGGGGCGCTGGATGATGAA | 2160 |
| Qy | 2161 | GAGAGGCGGAGCTGTCTCCCTCCGAGGTGCTATGCAACAGGCCATCCACCAATCGAG | 2220 |
| Db | 2161 | GAGAGGCGGAGCTGTCTCCCTCCGAGGTGCTATGCAACAGGCCATCCACCAATCGAG | 2220 |
| Qy | 2221 | TTCTGCTTGGGCTGGCTCTCCACACCGCTCTTACCTGCGCTGTGGGCGCTGAGGCTG | 2280 |
| Db | 2221 | TTCTGCTTGGGCTGGCTCTCCACACCGCTCTTACCTGCGCTGTGGGCGCTGAGGCTG | 2280 |
| Qy | 2281 | GCCACGCCACAGCTGTCCGAGGTTCTGTGGGCCATGTGTGATGCGCATAGGCGCTGGGCGTG | 2340 |
| Db | 2281 | GCCACGCCACAGCTGTCCGAGGTTCTGTGGGCCATGTGTGATGCGCATAGGCGCTGGGCGTG | 2340 |
| Qy | 2341 | GGCCGGAGGTTGGCGTGGGCGCTGTGTGTCTGTGTCTCCCATCTTTGCGGCTTTGCGGTG | 2400 |
| Db | 2341 | GGCCGGAGGTTGGCGTGGGCGCTGTGTGTCTGTGTCTCCCATCTTTGCGGCTTTGCGGTG | 2400 |

| | | | |
|----|------|---|------|
| Qy | 2401 | ATGACCGTGCTATCTCTGCTGGTGATGGAGGACTCTCAGCCTTCTCTGCACGCCCTCGG | 2460 |
| Db | 2401 | ATGACCGTGCTATCTCTGCTGGTGATGGAGGAGACTCTCAGCCTTCTCTGCACGCCCTCGG | 2460 |
| Qy | 2461 | CTGCACCTGGGTGGAAATTCACAGACCAAGTTCTACTCAGGCACGGGCTACAGCTGAGTCCC | 2520 |
| Db | 2461 | CTGCACCTGGGTGGAAATTCACAGAACAAAGTTCTACTCAGGCACGGGCTACAAAGCTGAGTCCC | 2520 |
| Qy | 2521 | TTCACTCTTCGCTGCCACAGATGACTAGGGGCCACTTGCAGGTCTCTGCCAGACCTCTCTTCCT | 2580 |
| Db | 2521 | TTCACTCTTCGCTGCCACAGATGACTAGGGGCCACTGCAGGTCTCTGCCAGACCTCTCTTCCT | 2580 |
| Qy | 2581 | GACCTCTGAGGCAGGAGAGGAATAAAGACGGTCCGCCCTCGCAAAAAAAAAAAAAAAAAA | 2640 |
| Db | 2581 | GACCTCTGAGGCAGGAGAGGAATAAAGACGGTCCGCCCTCGCAAAAAAAAAAAAAAAAAA | 2640 |

RESULT 3

```

US-09-016-434-1094
: Sequence 1094, Application US/09016434
: Patent No. 6500938
: GENERAL INFORMATION:
: APPLICANT: Janice Au-Young
: APPLICANT: Jeffrey J. Sailhamer
: TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
: TITLE OF INVENTION: PATHWAY GENE EXPRESSION
: NUMBER OF SEQUENCES: 1490
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
: STREET: 3174 PORTER DRIVE
: CITY: PALO ALTO
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/016,434
: FILING DATE: HEREWITH
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Zeller, Karen J.
: REGISTRATION NUMBER: 37,071
: REFERENCE/DOCKET NUMBER: PA-0002 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 855-0555
: TELEFAX: (650) 845-4166
: INFORMATION FOR SEQ ID NO: 1094:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2655 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GENBANK
: CLONE: g1245045
US-09-016-434-1094

```

```
Query Match      100.0%; Score 2640; DB 3; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

[illegible]

QY 61 GGCTCCATGTTCCGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCCCACAGCGGCT 120
DB 61 GGCTCCATGTTCCGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCCCACAGCGGCT 120
QY 121 GCCTACACCTGCGTAGTCCGGCTGGGCGAGCTGGGCTCTGTGAGTTTCAGAGACCTCAAC 180
DB 121 GCCTACACCTGCGTAGTCCGGCTGGGCGAGCTGGGCTCTGTGAGTTTCAGAGACCTCAAC 180
QY 181 GCCTCGGTAGCGCCCTTCCAGAGACGCTTTGTGTGATGTTTGGCGCTGTGAGGAGCTG 240
DB 181 GCCTCGGTAGCGCCCTTCCAGAGACGCTTTGTGTGATGTTTGGCGCTGTGAGGAGCTG 240
QY 241 GAGAAGACCTTCACTTCTGCAAGAGAGGAGGTGGGCGGCTGGGCTGGTCTGCCCCCG 300
DB 241 GAGAAGACCTTCACTTCTGCAAGAGAGGAGGTGGGCGGCTGGGCTGGTCTGCCCCCG 300
QY 301 CCAAGAGGAGGCTGCGGCACCCACACCCCGGACCTGTGCGCATCCAGGAGGAGACG 360
DB 301 CCAAGAGGAGGCTGCGGCACCCACACCCCGGACCTGTGCGCATCCAGGAGGAGACG 360
QY 361 GAGCGCTGCGCCAGGAGCTGCGGATGTGCGGGCAACACAGAGGCTTGCAGGCGCCAG 420
DB 361 GAGCGCTGCGCCAGGAGCTGCGGATGTGCGGGCAACACAGAGGCTTGCAGGCGCCAG 420
QY 421 CTGCACAGCTGAGCTTCCAGCGCGCGTGTCTAGCCAGGGCCATGAACTCAGCTGGCA 480
DB 421 CTGCACAGCTGAGCTTCCAGCGCGCGTGTCTAGCCAGGGCCATGAACTCAGCTGGCA 480
QY 481 GCGCGCCACACAGATGGGGCTCAGAGAGAGACCCCTGTCTCAGAGCCCGCGGGGCGG 540
DB 481 GCGCGCCACACAGATGGGGCTCAGAGAGAGACCCCTGTCTCAGAGCCCGCGGGGCGG 540
QY 541 CACAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGGTGGAGCCCCACAGGCCCCCTGCC 600
DB 541 CACAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGGTGGAGCCCCACAGGCCCCCTGCC 600
QY 601 CTAGAGCGCTGCTCTGAGGGCTGCGCGGCTTCTCTATGCGAGCTTCAGGAGCTG 660
DB 601 CTAGAGCGCTGCTCTGAGGGCTGCGCGGCTTCTCTATGCGAGCTTCAGGAGCTG 660
QY 661 GAGCAGCGCTGAGCAGATCCGACAGAGATCCGCAAGATCCGCAAGATCCAGGACTGC 720
DB 661 GAGCAGCGCTGAGCAGATCCGACAGAGATCCGCAAGATCCGCAAGATCCAGGACTGC 720
QY 721 TCTACTGGGTGAGCAGATCCGACAGAGATCCGCAAGATCCGCAAGATCCGCAAGATCC 780
DB 721 TCTACTGGGTGAGCAGATCCGACAGAGATCCGCAAGATCCGCAAGATCCAGGACTGC 780
QY 781 CAGCTCTTCCGTTCTGAGCAGGAGGAGGCGCGCTCGGGCCCTGCGAGCTGCAAC 840
DB 781 CAGCTCTTCCGTTCTGAGCAGGAGGAGGCGCGCTCGGGCCCTGCGAGCTGCAAC 840
QY 841 CAGCAGGCGAGGCTGAGGAGTCTCGGGGAGACAGAGCGGTTCTGAGCCAGGCTG 900
DB 841 CAGCAGGCGAGGCTGAGGAGTCTCGGGGAGACAGAGCGGTTCTGAGCCAGGCTG 900
QY 901 CTAGGCGGGTGTGAGCTGCTGCGCCAGGAGGAGTGCAGTCCAGAGATGAAGGCC 960
DB 901 CTAGGCGGGTGTGAGCTGCTGCGCCAGGAGGAGTGCAGTCCAGAGATGAAGGCC 960
QY 961 GTGTACTGGCCCTGAACTGAGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1020
DB 961 GTGTACTGGCCCTGAACTGAGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1020
QY 1021 GCTGTGTCTGTGAGACCTGCGCGCTGAGGAGGCTGCGGGAGCAGCTCGATG 1080
DB 1021 GCTGTGTCTGTGAGACCTGCGCGCTGAGGAGGCTGCGGGAGCAGCTCGATG 1080
QY 1081 GAGGAGGAGTGAAGTCCGTTGAGTCCAGGAGTCCAGGAGTCCAGGAGTCCAGGAGTCC 1140
DB 1081 GAGGAGGAGTGAAGTCCGTTGAGTCCAGGAGTCCAGGAGTCCAGGAGTCCAGGAGTCC 1140

QY 1141 ATCCGACCAACCGCTTACGGCCAGCTTCCAGGGCATCGTGGATCGCTACGGGTGGC 1200
DB 1141 ATCCGACCAACCGCTTACGGCCAGCTTCCAGGGCATCGTGGATCGCTACGGGTGGC 1200
QY 1201 CGCTACAGGAGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCCCTTGTGCT 1260
DB 1201 CGCTACAGGAGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCCCTTGTGCT 1260
QY 1261 GTGATGTTGGGGATGTGGSCCAAGGCTCTCATGTTCTTCTTGGCCCTGGCCATGGTC 1320
DB 1261 GTGATGTTGGGGATGTGGSCCAAGGCTCTCATGTTCTTCTTGGCCCTGGCCATGGTC 1320
QY 1321 CTTGGCAGAGACCGACCGGCTGTAAAGCCGCGCAGAACGAGATCTGGCAGACTTCTTC 1380
DB 1321 CTTGGCAGAGACCGACCGGCTGTAAAGCCGCGCAGAACGAGATCTGGCAGACTTCTTC 1380
QY 1381 AGGGGCGCTTACCTGCTCTGCTTATGGGCTGTCTTCACTACACCGGCTTCACTAC 1440
DB 1381 AGGGGCGCTTACCTGCTCTGCTTATGGGCTGTCTTCACTACACCGGCTTCACTAC 1440
QY 1441 AACGAGTGTTCAGTCCGCGCACAGCATCTTCCCTCGGGCTGGAGTGTGGCGCCATG 1500
DB 1441 AACGAGTGTTCAGTCCGCGCACAGCATCTTCCCTCGGGCTGGAGTGTGGCGCCATG 1500
QY 1501 GCCAACAGTCTGGCTGGAGTGTATGATTTCTGGCCCGCAGCACAGATGCTTACCTGGAT 1560
DB 1501 GCCAACAGTCTGGCTGGAGTGTATGATTTCTGGCCCGCAGCACAGATGCTTACCTGGAT 1560
QY 1561 CCCAACAGTCAACCGCTTCTTCTGGGACCTACCTTGGGATCGATCTTATTTGGAGC 1620
DB 1561 CCCAACAGTCAACCGCTTCTTCTGGGACCTACCTTGGGATCGATCTTATTTGGAGC 1620
QY 1621 CTGGCTGCCAACACCTTGGCTTCTCAACTCTTCAAGATGAAGATGCTCCGTCATCTG 1680
DB 1621 CTGGCTGCCAACACCTTGGCTTCTCAACTCTTCAAGATGAAGATGCTCCGTCATCTG 1680
QY 1681 GGGCTGTCGACATGCGCTTGGGGTGGTCTCGGAGCTTCAACACAGCTGCACTTTGGC 1740
DB 1681 GGGCTGTCGACATGCGCTTGGGGTGGTCTCGGAGCTTCAACACAGCTGCACTTTGGC 1740
QY 1741 CAGAGCACCGGCTGCTGGAGAGCTGCGGAGCTCACCTTCTGCTGGGACTCTTC 1800
DB 1741 CAGAGCACCGGCTGCTGGAGAGCTGCGGAGCTCACCTTCTGCTGGGACTCTTC 1800
QY 1801 GGTACCTCGTGTCTTAGTTCATCAAGTGGCTGTGTCTGGGCTGCGAGGCGGCC 1860
DB 1801 GGTACCTCGTGTCTTAGTTCATCAAGTGGCTGTGTCTGGGCTGCGAGGCGGCC 1860
QY 1861 TCGCCAGCATCTCTATCCACTTTCATCAACATGTTCTTCTTCCACAGCCCGCAGAAC 1920
DB 1861 TCGCCAGCATCTCTATCCACTTTCATCAACATGTTCTTCTTCCACAGCCCGCAGAAC 1920
QY 1921 AGGCTGCTTACCCCGGAGGAGTGGTCCAGGCGCACGCTGTGTCTTGGGCTTGGCC 1980
DB 1921 AGGCTGCTTACCCCGGAGGAGTGGTCCAGGCGCACGCTGTGTGTCTTGGGCTTGGCC 1980
QY 1981 ATGTTGCCCATCTCTGCTGTGGCACAACCCCTGCACTGCTGCAACCGCCCGCC 2040
DB 1981 ATGTTGCCCATCTCTGCTGTGGCACAACCCCTGCACTGCTGCAACCGCCCGCC 2040
QY 2041 CTTGGAGGAGGCGCGCTGACCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100
DB 2041 CTTGGAGGAGGCGCGCTGACCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100
QY 2101 GACGATCTGTAATGGCTGGAGTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
DB 2101 GACGATCTGTAATGGCTGGAGTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
QY 2161 GAGGAGGCGGAGCTGCTCCCTCCGAGGTGCTCATGACAGGAGGAGGAGGAGGAGGAG 2220
DB 2161 GAGGAGGCGGAGCTGCTCCCTCCGAGGTGCTCATGACAGGAGGAGGAGGAGGAGGAG 2220
QY 2221 TTTCTGCTGGGCTGGGCTTCCAAACCCGCTCTTCTTCTGCGGCTGTGGGCTTGAAGCTG 2280

| | | | |
|----|------|---|------|
| QY | 61 | GGCTCCATGTTCCGGAGCGAGGAGTGGCCCTGGTCCAGCTCTTTCTGCCACAGCGGCT | 120 |
| Db | 61 | GGCTCCATGTTCCGGAGCGAGGAGTGGCCCTGGTCCAGCTCTTTCTGCCACAGCGGCT | 120 |
| QY | 121 | GGCTACACCTGCGTAGTGGCTGGCGAGCTGGGCTCGTGGAGTTACAGACCTCAAC | 180 |
| Db | 121 | GGCTACACCTGCGTAGTGGCTGGCGAGCTGGGCTCGTGGAGTTACAGACCTCAAC | 180 |
| QY | 181 | GGCTCGGTGAGCGCTTCAGAGACGCTTTGTGGTTGATGTTTGGCGCTGTGAGGAGCTG | 240 |
| Db | 181 | GGCTCGGTGAGCGCTTCAGAGACGCTTTGTGGTTGATGTTTGGCGCTGTGAGGAGCTG | 240 |
| QY | 241 | GAGAGACCTTCACCTTCCTGTCAGAGAGAGGTGGCGGCTGGGCTGGTCTGCCCGG | 300 |
| Db | 241 | GAGAGACCTTCACCTTCCTGTCAGAGAGAGGTGGCGGCTGGGCTGGTCTGCCCGG | 300 |
| QY | 301 | CCAAAGGGAGGTGCGCGACCCCCACCCCGGACCTGCTGCGATCCAGAGAGAGAG | 360 |
| Db | 301 | CCAAAGGGAGGTGCGCGACCCCCACCCCGGACCTGCTGCGATCCAGAGAGAGAG | 360 |
| QY | 361 | GAGCGCTGGCCAGAGAGTGGGGATGTGGGGGCAACAGAGGCTCTGGCGGCCAG | 420 |
| Db | 361 | GAGCGCTGGCCAGAGAGTGGGGATGTGGGGGCAACAGAGGCTCTGGCGGCCAG | 420 |
| QY | 421 | CTGCACACGCTGCAGCTCCAGCGCGCTGTACGCCAGGGCCATGAACCTCAGCTGGCA | 480 |
| Db | 421 | CTGCACACGCTGCAGCTCCAGCGCGCTGTACGCCAGGGCCATGAACCTCAGCTGGCA | 480 |
| QY | 481 | GCGGCCACACAGATGGGGCTCAGAGAGAGCGCCCTGCTCCAGGCCCCCGGGGGCG | 540 |
| Db | 481 | GCGGCCACACAGATGGGGCTCAGAGAGAGCGCCCTGCTCCAGGCCCCCGGGGGCG | 540 |
| QY | 541 | CACGAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGGTGGAGGCCCAAGGCCCTGGC | 600 |
| Db | 541 | CACGAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGGTGGAGGCCCAAGGCCCTGGC | 600 |
| QY | 601 | CTAGAGCGCTGCTCTGAGAGGCTGCGCGGCTTCTCATTTGCCAGCTTCAGGAGCTG | 660 |
| Db | 601 | CTAGAGCGCTGCTCTGAGAGGCTGCGCGGCTTCTCATTTGCCAGCTTCAGGAGCTG | 660 |
| QY | 661 | GAGCAGCGCTGAGACACCCGCTGACGGCGAGCCAGCCAGCTGGATGACCTTCTCATC | 720 |
| Db | 661 | GAGCAGCGCTGAGACACCCGCTGACGGCGAGCCAGCCAGCTGGATGACCTTCTCATC | 720 |
| QY | 721 | TCTTACTGGGTGAGCAGATCGACAGAGATCCGCAAGATCACGGACTGCTTCACTGC | 780 |
| Db | 721 | TCTTACTGGGTGAGCAGATCGACAGAGATCCGCAAGATCACGGACTGCTTCACTGC | 780 |
| QY | 781 | CACGTCTTCCGTTCTGACAGGAGGAGGCGCGCTCGGGGCTCGAGCAGCTGCA | 840 |
| Db | 781 | CACGTCTTCCGTTCTGACAGGAGGAGGCGCGCTCGGGGCTCGAGCAGCTGCA | 840 |
| QY | 841 | CAGCAGCAGCAGAGCTGAGAGGTCTCTCGGGAGACAGAGCGGTCTCTGAGCCAGGTG | 900 |
| Db | 841 | CAGCAGCAGCAGAGCTGAGAGGTCTCTCGGGAGACAGAGCGGTCTCTGAGCCAGGTG | 900 |
| QY | 901 | CTAGGCGGGTCTGAGCTGCTGCGCCAGGGCAGGTCCAGATGAAGAGGCC | 960 |
| Db | 901 | CTAGGCGGGTCTGAGCTGCTGCGCCAGGGCAGGTCCAGATGAAGAGGCC | 960 |
| QY | 961 | GTGTACTGGCTGAAACAGTGCAGCTGAGCAGCAGCAGCTGCGGCTCATTTGCCGAG | 1020 |
| Db | 961 | GTGTACTGGCTGAAACAGTGCAGCTGAGCAGCAGCAGCTGCGGCTCATTTGCCGAG | 1020 |
| QY | 1021 | GGCTGGTCTGTGCGAGACCTGGCCCTCTGAGAGGCGCTTGGGAGAGCTCGATG | 1080 |
| Db | 1021 | GGCTGGTCTGTGCGAGACCTGGCCCTCTGAGAGGCGCTTGGGAGAGCTCGATG | 1080 |
| QY | 1081 | GAGGAGGAGTGAATGCTGAGTCAACCGATCCCTGCGGGACATGCCCCCAGCTC | 1140 |
| Db | 1081 | GAGGAGGAGTGAATGCTGAGTCAACCGATCCCTGCGGGACATGCCCCCAGCTC | 1140 |
| QY | 1141 | ATCCGACCAACCGCTTCAAGCCAGCTTCAAGGCGATCGTGATCGCTACGCGTGGC | 1200 |
| Db | 1141 | ATCCGACCAACCGCTTCAAGCCAGCTTCAAGGCGATCGTGATCGCTACGCGTGGC | 1200 |
| QY | 1201 | CGCTACACGAGGTCAACCCCGCTTCAACCATCATCACCTTCCCTTCTGTTGCT | 1260 |
| Db | 1201 | CGCTACACGAGGTCAACCCCGCTTCAACCATCATCACCTTCCCTTCTGTTGCT | 1260 |
| QY | 1261 | GTGATGTTGGGGATGTGGGCAAGGCTGTCTTCTTCCGCTGGCCATGGTC | 1320 |
| Db | 1261 | GTGATGTTGGGGATGTGGGCAAGGCTGTCTTCTTCCGCTGGCCATGGTC | 1320 |
| QY | 1321 | CTTGGCGAGAACCGACCGCTGTGAAGCGCGGAGAACGAGATCTGGCAGATTTCTTC | 1380 |
| Db | 1321 | CTTGGCGAGAACCGACCGCTGTGAAGCGCGGAGAACGAGATCTGGCAGATTTCTTC | 1380 |
| QY | 1381 | AGGGGCGCTACTCTCTCTGTTATGGGCTGTCTCATCTACACCGGCTTCATCTAC | 1440 |
| Db | 1381 | AGGGGCGCTACTCTCTCTGTTATGGGCTGTCTCATCTACACCGGCTTCATCTAC | 1440 |
| QY | 1441 | AACGATGCTTCAAGTGGGCGCACAGATCTTCCCTCGGGCTGAGTGTGGCGCGCATG | 1500 |
| Db | 1441 | AACGATGCTTCAAGTGGGCGCACAGATCTTCCCTCGGGCTGAGTGTGGCGCGCATG | 1500 |
| QY | 1501 | GCCAAACAGTCTGGGTGGAGTGAATCTTCTGGGCGCACAGATCTTACCTGGAT | 1560 |
| Db | 1501 | GCCAAACAGTCTGGGTGGAGTGAATCTTCTGGGCGCACAGATCTTACCTGGAT | 1560 |
| QY | 1561 | CCCAAAGTCAAGGCTGCTTCTGGGACCTTACCTTGGCATCGATCTTATTTGGAGC | 1620 |
| Db | 1561 | CCCAAAGTCAAGGCTGCTTCTGGGACCTTACCTTGGCATCGATCTTATTTGGAGC | 1620 |
| QY | 1621 | CTGGCTGCGAACCACTTGAAGTCTTCAATCTCTTCAAGATGAAGATGTCGTCATCTG | 1680 |
| Db | 1621 | CTGGCTGCGAACCACTTGAAGTCTTCAATCTCTTCAAGATGAAGATGTCGTCATCTG | 1680 |
| QY | 1681 | GCGCTGCTGACATGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGG | 1740 |
| Db | 1681 | GCGCTGCTGACATGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGG | 1740 |
| QY | 1741 | CAGAGCAGCGCTGCTGAGAGCGCTGCGGAGCTCACCTTCTCTGCTGGGACTCTTC | 1800 |
| Db | 1741 | CAGAGCAGCGCTGCTGAGAGCGCTGCGGAGCTCACCTTCTCTGCTGGGACTCTTC | 1800 |
| QY | 1801 | GGTTACTGCTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 1860 |
| Db | 1801 | GGTTACTGCTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 1860 |
| QY | 1861 | TGCGCCAGCATCTCATCTTCAATCAATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 1920 |
| Db | 1861 | TGCGCCAGCATCTCATCTTCAATCAATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 1920 |
| QY | 1921 | AGGTGCTCTTACCCCGCAGAGGTGTTCAGGCGACGCTGGTGGTCTTGGCTTGGCTTGGCT | 1980 |
| Db | 1921 | AGGTGCTCTTACCCCGCAGAGGTGTTCAGGCGACGCTGGTGGTCTTGGCTTGGCTTGGCT | 1980 |
| QY | 1981 | ATGTTGCCATCTGCTGCTTGGGACACCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 2040 |
| Db | 1981 | ATGTTGCCATCTGCTGCTTGGGACACCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT | 2040 |
| QY | 2041 | CTGCGGAGGAGGCGCGCTGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG | 2100 |
| Db | 2041 | CTGCGGAGGAGGCGCGCTGACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG | 2100 |
| QY | 2101 | GACGATCTGTGAATGGCTGGAGCTTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG | 2160 |
| Db | 2101 | GACGATCTGTGAATGGCTGGAGCTTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG | 2160 |
| QY | 2161 | GAG | 2220 |
| Db | 2161 | GAG | 2220 |
| QY | 2221 | TTTCTGCTGGGCTGCTGCTTCAACACCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT | 2280 |

1381 GCAGAACGAGATCTGGCAGACTTTCTTTCAGGGGCGCTACCTGCTCTCTTATGGGCT 1440
1413 GTTCTCCATCTACACGGGCTTCACTACAAAGAGTGCTTCACTGCGGCCACACGATCTT 1472
1441 GTTCTCCATCTACACGGGCTTCACTACAAAGAGTGCTTCACTGCGGCCACACGATCTT 1500
1473 CCCTCGGGCTGGAGTGTGGCGCCATGGCCAAACAGTCTGGCTGGAGTGATGCAATCT 1532
1501 CCCTCGGGCTGGAGTGTGGCGCCATGGCCAAACAGTCTGGCTGGAGTGATGCAATCT 1560
1533 GCGCCAGCACAAGATCTTACCTTGGATCCCAACGTCACCGGTGCTTCTTGGGACCTTA 1592
1561 GCGCCAGCACAAGATCTTACCTTGGATCCCAACGTCACCGGTGCTTCTTGGGACCTTA 1620
1593 CCGCTTTGGCATCGATCTTATTTGGAGCTTGGCTGCAACCACTTGGAGTCTTCTCAATC 1652
1621 CCGCTTTGGCATCGATCTTATTTGGAGCTTGGCTGCAACCACTTGGAGTCTTCTCAATC 1680
1653 CTTCAAGATGAAGATGTCGCTCATCTGGGCGTCTGTCACATGGCCCTTTGGGCTGCTCT 1712
1681 CTTCAAGATGAAGATGTCGCTCATCTGGGCGTCTGTCACATGGCCCTTTGGGCTGCTCT 1740
1713 CGAGTCTTCAACCAAGTGCATCTTGGCCAGAGGCCACCGGCTGCTGTCGAGACGCTGCC 1772
1741 CGAGTCTTCAACCAAGTGCATCTTGGCCAGAGGCCACCGGCTGCTGTCGAGACGCTGCC 1800
1773 GAGCTCACCTTCTGCTGGAGCTTCTCGTTACCTGCTGCTGCTTCTAGTCACTACAGTG 1832
1801 GAGCTCACCTTCTGCTGGAGCTTCTCGTTACCTGCTGCTGCTTCTAGTCACTACAGTG 1860
1833 GCTGTGTGCTGGCTGTCAGAGCGCGCTCG--CCAGCATCTTCACTCACTTCACTAA 1889
1861 GCTGTGTGCTGGCTGTCAGAGCGCGCTCGCGCCGCCATCTCTATCTCACTTCACTAA 1920
1890 CATGTTCTCTTCTCCACAGCCCCAGCAACAGGCTGCTTACCCCGGACGAGGTGCT 1949
1921 CATGTTCTCTTCTCCACAGCCCCAGCAACAGGCTGCTTACCCCGGACGAGGTGCT 1980
1950 CAGGCGACGCTGGTGGTCTTGGCTTGGCCATGCTGTCCTGCTGCTGCTGGCCACAC 2009
1981 CAGGCGACGCTGGTGGTCTTGGCTTGGCCATGCTGTCCTGCTGCTGCTGGCCACAC 2040
2010 CTTGACCTGCTGACCGCGCACCGCGCTGCGGAGGCGCGCTGACCGACAGCA 2069
2041 CTTGACCTGCTGACCGCGCACCGCGCTGCGGAGGCGCGCTGACCGACAGCA 2100
2070 GGAACAAAGCGGGTGTCTGGAATCTGCTGACCGCATCTGTGAATGGCTGGAGCTCGA 2129
2101 GGAACAAAGCGGGTGTCTGGAATCTGCTGACCGCATCTGTGAATGGCTGGAGCTCGA 2160
2130 TGAGGAAAGGCGGGGCTTGATGATGAAGAGGCGGAGCTGCTGCTGCTGCTGCTGCT 2189
2161 TGAGGAAAGGCGGGGCTTGATGATGAAGAGGCGGAGCTGCTGCTGCTGCTGCTGCT 2220
2190 GCTCATGACACAGGCCATCCACACATCGAGTTCTGCTGGGCTGCTGCTGCTGCTGCT 2249
2221 GCTCATGACACAGGCCATCCACACATCGAGTTCTGCTGGGCTGCTGCTGCTGCTGCT 2280
2250 CTTCTACCTGCGCTTGTGGGCTTGAAGCTGCGCCCAACCGCTGCTGCTGCTGCTGCT 2309
2281 CTTCTACCTGCGCTTGTGGGCTTGAAGCTGCGCCCAACCGCTGCTGCTGCTGCTGCT 2340
2310 GGCCTATGATGCGCATAGGCTTGGCTGCGGCGGAGGCTGGGCTGCGGCTGCTGCT 2369
2341 GGCCTATGATGCGCATAGGCTTGGCTGCGGCGGAGGCTGGGCTGCGGCTGCTGCT 2400
2370 GCTGTGCTCCCATCTTTGCGGCTTGTGCGTGTATGACCGGTGCTATCTGCTGCTGATGA 2429
2401 GCTGTGCTCCCATCTTTGCGGCTTGTGCGTGTATGACCGGTGCTATCTGCTGCTGATGA 2460
2430 GGGACTCTCAGCTTCTCTGCA CGCCCTGCGGCTGCACTGGGTGGAATTCAGAAACAAGTT 2489

2461 GGGACTCTCAGCTTCTCTGCAACGCTGCGGCTGCACTGGGTGGAATTCAGAAACAAGTT 2520
2490 CTACTCAGGACGGGCTACAGCTGAGTCCCTTCACTTCTGCTGCCACAGATGACTAGGG 2549
2521 CTACTCAGGACGGGCTACAGCTGAGTCCCTTCACTTCTGCTGCCACAGATGACTAGGG 2580
2550 CCCTCGAGGCTCTGCGGACGCTTCTTCTGACCTCTGAGGACGAGAGGAATAAAGAC 2609
2581 CCCTCGAGGCTCTGCGGACGCTTCTTCTGACCTCTGAGGACGAGAGGAATAAAGAC 2640
2610 GGTCCGCCCTGGCA 2623
2641 GGTCCGCCCTGGCA 2654

RESULT 7

US-09-949-016-2405
; Sequence 2405, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 2405
; LENGTH: 2654
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2405

Query Match 97.7%; Score 2580; DB 3; Length 2654;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2585; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

33 GCGCGCGCAGACACACCGCGGACCATGGGCTTCCATGTTCCGGAGGAGGAGGTGGCCCT 92
61 GCGGCGCGCAGACACACCGCGGACCATGGGCTTCCATGTTCCGGAGGAGGAGGTGGCCCT 120
93 GGTCCAGCTCTTTCTGCGCCACAGCGGCTGCTTACCTCGGTGAGTGGGCTGGGCGAGCT 152
121 GGTCCAGCTCTTTCTGCGCCACAGCGGCTGCTTACCTCGGTGAGTGGGCTGGGCGAGCT 180
153 GGGCTCTGTGAGTTCAGAGACCTCAACGCTCTGGTGAGCGCTTTCAGAGACGCTTTGT 212
181 GGGCTCTGTGAGTTCAGAGACCTCAACGCTCTGGTGAGCGCTTTCAGAGACGCTTTGT 240
213 GGTTCATGTTGGCGCTGTGAGGAGCTGAGAGACCTTCACTTCTCTGTCAGAGAGGT 272
241 GGTTCATGTTGGCGCTGTGAGGAGCTGAGAGACCTTCACTTCTCTGTCAGAGAGGT 300
273 GCGGCGGCTGGGCTGGTCTTCCCGCCCGGAGGAGGCTGCGGACACCCCGCCCG 332
301 GCGGCGGCTGGGCTGGTCTTCCCGCCCGGAGGAGGCTGCGGACACCCCGCCCG 360
333 GGAATCTGTGCGCATCCAGAGGAGACGAGAGCGCTTGGCCCAAGAGCTGCGGATGTGCG 392
361 GGAATCTGTGCGCATCCAGAGGAGACGAGAGCGCTTGGCCCAAGAGCTGCGGATGTGCG 420
393 GGGCAACAGAGGCGCTTGGGCGGCTGAGTCACTGAGTCCAGCGCGGCTGCT 452
421 GGGCAACAGAGGCGCTTGGGCGGCTGAGTCACTGAGTCCAGCGCGGCTGCT 480
453 AGCCAGGGCCATGAACCTCAGCTGGCAGCGCCCAACAGATGGGCGCTCAGAGAGGAC 512

[illegible]

| | | |
|------|---|------|
| 1561 | GGCCACGACACGATGCTTACCTCTGGATCCAAACGTCAACGGTGTCTTCTCTGGGACCCCTA | 1623 |
| 1593 | CCCTTTGGCATCGATCCTATTATGGAGCCCTGGCTGCGCAACACTTGGACTTCCTCAACTC | 1655 |
| 1621 | CCCTTTGGCATCGATCCTATTATGGAGCCCTGGCTGCGCAACACTTGGACTTCCTCAACTC | 1680 |
| 1653 | CTTCAAGATGAAGATGTCCGTCATCTCTGGGCGTCTGCGCATGSGGCTTTTGGGGTGGTCCCT | 1712 |
| 1681 | CTTCAAGATGAAGATGTCCGTCATCTCTGGGCGTCTGCGCATGSGGCTTTTGGGGTGGTCCCT | 1740 |
| 1713 | CGGAGTCTTCAACACAGTGCACCTTTGGCCAGAGGACCGGCTGCTGCTGGAGAGACGCTGCC | 1772 |
| 1741 | CGGAGTCTTCAACACAGTGCACCTTTGGCCAGAGGACCGGCTGCTGCTGGAGAGACGCTGCC | 1800 |
| 1773 | GGAGTCAACCTTCTGCTGGGACTCTTCGGTTACCTCTGCTTCCCTAGTCACTACAAAGTG | 1832 |
| 1801 | GGAGTCAACCTTCTGCTGGGACTCTTCGGTTACCTCTGCTTCCCTAGTCACTACAAAGTG | 1860 |
| 1833 | GCTGTGTGTCTGGGCTGCCAGGGCCGCTCG---CCAGCATCTCTCATCTTCATCAACAA | 1889 |
| 1861 | GCTGTGTGTCTGGGCTGCCAGGGCCGCTCGGGCCCCAGCATCTCTCATCTTCATCAACAA | 1920 |
| 1890 | CATGTTCTCTTCTTCCACAGCCCGCAGCAACAGGCTGCTCTACCCCGGCAAGAGTGGT | 1949 |
| 1921 | CATGTTCTCTTCTTCCACAGCCCGCAGCAACAGGCTGCTCTACCCCGGCAAGAGTGGT | 1980 |
| 1950 | CCAGGCCACGCTGCTGTGCTCTGGCCTTTGGCCATGTGGCCATCTGCTGCTTTGGCACACC | 2009 |
| 1981 | CCAGGCCACGCTGCTGTGCTCTGGCCTTTGGCCATGTGGCCATCTGCTGCTTTGGCACACC | 2040 |
| 2010 | CCTGCACCTGTGCACCGCCACCGCCCGCTCTGGGAGAGGCCCGCTGCACCGACAGGA | 2069 |
| 2041 | CCTGCACCTGTGCACCGCCACCGCCCGCTCTGGGAGAGGCCCGCTGCACCGACAGGA | 2100 |
| 2070 | GGAAAAAAGCCCGGGTTGCTGACACTGCTGACCGATCTGTGAATGGCTGGAGCTCGGA | 2129 |
| 2101 | GGAAAAAAGCCCGGGTTGCTGACACTGCTGACCGATCTGTGAATGGCTGGAGCTCGGA | 2160 |
| 2130 | TCAGGAAAGACGAGGGGCTGTGATGATGAAGAGAGGCGCAGAGCTGCTCCCTCCGAGGT | 2189 |
| 2161 | TGAGGAAAGACGAGGGGCTGTGATGATGAAGAGAGGCGCAGAGCTGCTCCCTCCGAGGT | 2220 |
| 2190 | GCTCATGCAACAGGCCATCCACACCATCGAGTTCTGCCCTGGGCTGCGTCTCCAAACCCGC | 2249 |
| 2221 | GCTCATGCAACAGGCCATCCACACCATCGAGTTCTGCCCTGGGCTGCGTCTCCAAACCCGC | 2280 |
| 2250 | CTTCTACTGGCTGTGGGCCCTGAGCCTGGGCCACCGCCAGCTGTCCGAGGTTCTGTGT | 2309 |
| 2281 | CTTCTACTGGCTGTGGGCCCTGAGCCTGGGCCACCGCCAGCTGTCCGAGGTTCTGTGT | 2340 |
| 2310 | GGCCATGATGATGGCATAGGCTGTGGGCTGGGCGGAGAGGTGGGCGTGGCGCTGTGGT | 2369 |
| 2341 | GGCCATGATGATGGCATAGGCTGTGGGCTGGGCGGAGAGGTGGGCGTGGCGCTGTGGT | 2400 |
| 2370 | GCTGTGCTCCCATCTTTTGGCCGCTTTGCCCGTGAAGCCGTGCTATCTGCTGGTGATGGA | 2429 |
| 2401 | GCTGTGCTCCCATCTTTTGGCCGCTTTGCCCGTGAAGCCGTGCTATCTGCTGGTGATGGA | 2460 |
| 2430 | GGGACTCTCAGCTTCTTGCAACGCCCTGGGCTGCACTGGGTTGGAAATTCAGAAACAAGTT | 2489 |
| 2461 | GGGACTCTCAGCTTCTTGCAACGCCCTGGGCTGCACTGGGTTGGAAATTCAGAAACAAGTT | 2520 |
| 2490 | CTPACTCAGGCACGGGCTACAGCTGAGTCCCTTTCACCTTCGCTGCGCAACAGATGACTAGGG | 2549 |
| 2521 | CTPACTCAGGCACGGGCTACAGCTGAGTCCCTTTCACCTTCGCTGCGCAACAGATGACTAGGG | 2580 |
| 2550 | CCCACTCAGGTCCTGCGCAGACCTCTTCTGACCTCTCAGGCAGGAGAGGAATAAGAC | 2609 |
| 2581 | CCCACTCAGGTCCTGCGCAGACCTCTTCTGACCTCTCAGGCAGGAGAGGAATAAGAC | 2640 |
| 2610 | GGTTCGCGCCTGGCA | 2623 |
| 2641 | GGTTCGCGCCTGGCA | 2654 |

1780 ACCTTCTGCTGGAGCTCTTCGGTTACTCTGTTTCTAGTATCTCAAGTGTGCTGT 1839
1860 ACCTTCTGCTGGAGCTCTTCGGTTACTCTGTTTCTAGTATCTCAAGTGTGCTGT 1919
1840 GTCTGGGCTGCAAGGCGGCTCG--CCAGCATCTCTATCCACTTCAACATGTC 1896
1920 GTCTGGGCTGCAAGGCGGCTCGGCCCCAGCATCTCTATCCACTTCAACATGTC 1979
1897 CTCTTCTCCACAGCCCCAGCAACAGAGCTGCTTACCCCCGCGAGGAGTGTCCAGGCC 1956
1980 CTCTTCTCCACAGCCCCAGCAACAGAGCTGCTTACCCCCGCGAGGAGTGTCCAGGCC 2039
1957 ACCTGCTGCTGCTGGGCTTGGCCATGTTGCCATCTGCTGCTTGGCAGACCCCTGCAC 2016
2040 ACGTGTGCTGCTGGGCTTGGCCATGTTGCCATCTGCTGCTTGGCAGACCCCTGCAC 2099
2017 CTGCTGACCCGCGCCGCGCTCGGAGAGGCGCGCTGACCCAGGAGGAGAAAC 2076
2100 CTGCTGACCCGCGCCGCGCTCGGAGAGGCGCGCTGACCCAGGAGGAGAAAC 2159
2077 AAGCGCGGTTGCTGGAACCTGCTGACGCAATCTGTGAATGCTGGAAGTCCGATGAGAA 2136
2160 AAGCGCGGTTGCTGGAACCTGCTGACGCAATCTGTGAATGCTGGAAGTCCGATGAGAA 2219
2137 AAGCGAGGCGCTGGAATGATGAAGAGGAGCGAGCTGCTCCCTCCGAGGTGCTCATG 2196
2220 AAGCGAGGCGCTGGAATGATGAAGAGGAGCGAGCTGCTCCCTCCGAGGTGCTCATG 2279
2197 CACAGGCAATCACACCATCTGAGTCTGCTGCGCTGCTGCTCCACACCGCTCTCAT 2256
2280 CACAGGCAATCACACCATCTGAGTCTGCTGCGCTGCTGCTCCACACCGCTCTCAT 2339
2257 CTGCGCTGCTGGGCTGCTGAGCTGCGCCACCGCCAGCTGCTCGAGGTTCTGTGGGCCATG 2316
2340 CTGCGCTGCTGGGCTGCTGAGCTGCGCCACCGCCAGCTGCTCGAGGTTCTGTGGGCCATG 2399
2317 GTGATGCGCATAGGCTTGGGCTTGGGCGGAGGAGTGGGCGTGTGCTGCTGCTGCT 2376
2400 GTGATGCGCATAGGCTTGGGCTTGGGCGGAGGAGTGGGCGTGTGCTGCTGCTGCT 2459
2377 CCATCTTTGCGGCTTGTGCGGTGATGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2436
2460 CCATCTTTGCGGCTTGTGCGGTGATGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2519
2437 TCAGCTTCTGCAAGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2496
2520 TCAGCTTCTGCAAGCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2579
2497 GGCAAGGCTTACAGCTGCTGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 2556
2580 GGCAAGGCTTACAGCTGCTGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 2639
2557 CAGGCTCTGCAAGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 2616
2640 CAGGCTCTGCAAGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCT 2699
2617 CTTGGCA 2623
2700 CTTGGCA 2706

RESULT 9

US-09-949-016-2976
; Sequence 2976, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2976
; LENGTH: 2706
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-2976

Query Match 96.7%; Score 2553.4; DB 3; Length 2706;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2573; Conservative 0; Mismatches 11; Indels 3; Gaps 1;
QY 40 GCAGCACACCGGGGACCATGGGCTCCATGTTCCGAGCGAGGAGTGGCCCTGCTCCAG 99
Db 120 GGAGCCAGCCACAGGACCATGGGCTCCATGTTCCGAGCGAGGAGTGGCCCTGCTCCAG 179
QY 100 CTCTTTCTGCCACAGCGGCTGCTACCTGCTGAGTGGCTGGCGAGCTGGGCTC 159
Db 180 CTCTTTCTGCCACAGCGGCTGCTACCTGCTGAGTGGCTGGCGAGCTGGGCTC 239
QY 160 GTGAGTTTACAGACCTCAACGCTCGGTGAGCGCTTCCAGAGACGCTTTGTGTTGAT 219
Db 240 GTGAGTTTACAGACCTCAACGCTCGGTGAGCGCTTCCAGAGACGCTTTGTGTTGAT 299
QY 220 GTTTGGGCTGTGAGGAGCTGGAGAGACCTTCACTTCTTCCAGAGGAGTGGCGGG 279
Db 300 GTTTCGGGCTGTGAGGAGCTGGAGAGACCTTCACTTCTTCCAGAGGAGTGGCGGG 359
QY 280 GCTGGGCTGTCTGCCCCCGCAAGGGGAGGCTGCGGACCCCGGAGCTG 339
Db 360 GCTGGGCTGTCTGCCCCCGCAAGGGGAGGCTGCGGACCCCGGAGCTG 419
QY 340 CTCGCGATCAGAGGAGAGCGAGCGGCTGCGCCCGAGGAGTGGGGGCAAC 399
Db 420 CTCGCGATCAGAGGAGAGCGAGCGGCTGCGCCCGAGGAGTGGCGGGCAAC 479
QY 400 CAGCAGGCTTGGGGGCGGAGCTGACAGCTGAGCTTCAAGCGGCTGCTACGCGAG 459
Db 480 CAGCAGGCTTGGGGGCGGAGCTGACAGCTGAGCTTCAAGCGGCTGCTACGCGAG 539
QY 460 GGCATGAACTCAGCTGGCAGCGCGCCACACAGATGGGGCTCAGAGAGAGCGCCCTG 519
Db 540 GGCATGAACTCAGCTGGCAGCGCGCCACACAGATGGGGCTCAGAGAGAGCGCCCTG 599
QY 520 CTCAGGCGGCGGGGGCGGACCAAGGAGCTTCAAGGCTTCACTTTGTGGAGTGGCGGTG 579
Db 600 CTCAGGCGGCGGGGGCGGACCAAGGAGCTTCAAGGCTTCACTTTGTGGAGTGGCGGTG 659
QY 580 GAGCGGCGGCGGCGGCTGAGCGGCTGCTGAGGAGGCTGCGCGGCTTCTC 639
Db 660 GAGCGGCGGCGGCGGCTGAGCGGCTGCTGAGGAGGCTGCGCGGCTTCTC 719
QY 640 ATTGCGAGCTTCAAGGAGCTGGAGAGCGGCTGAGCAGCCCGCTGACGCGGAGCGAGCC 699
Db 720 ATTGCGAGCTTCAAGGAGCTGGAGAGCGGCTGAGCAGCCCGCTGACGCGGAGCGAGCC 779
QY 700 ACGTGGATGACCTTCTCTCATCTCTCTGAGGCTGAGCAGATCGGACAGAGATCCGCAAG 759
Db 780 ACGTGGATGACCTTCTCTCATCTCTCTGAGGCTGAGCAGATCGGACAGAGATCCGCAAG 839
QY 760 ATCAGGAGCTGCTTCACTGCTGCTTCTTCTGAGGAGGAGGAGGCGGCTC 819
Db 840 ATCAGGAGCTGCTTCACTGCTGCTTCTTCTGAGGAGGAGGAGGCGGCTC 899
QY 820 GGGGCGGCTGAGCAGCTGCAACAGCAGAGCGGAGCTGAGGAGGCTTCTCGGGAGACA 879
Db 900 GGGGCGGCTGAGCAGCTGCAACAGCAGAGCGGAGCTGAGGAGGCTTCTCGGGAGACA 959

```
QY 880 GAGCGGTTCTGAGCCAGGTGCTAGGCCGGGTGCTGAGCTGCTGCGCCGCGCAGGCGAGTG 939
Db 960 GAGCGGTTCTGAGCCAGGTGCTAGGCCGGGTGCTGAGCTGCTGCGCCGCGCAGGCGAGTG 1019
QY 940 CAGGTCCACAAGATGAAGCCCGTGTACTTGGCCCTGAACCAAGTGAAGCGTGAAGCAACG 999
Db 1020 CAGGTCCACAAGATGAAGCCCGTGTACTTGGCCCTGAACCAAGTGAAGCGTGAAGCAACG 1079
QY 1000 CACAAGTGCCTCATTTGCCAGGCTGCTGCTGCTGAGAGAGCTGCGCCGCTGCGAGAG 1059
Db 1080 CACAAGTGCCTCATTTGCCAGGCTGCTGCTGCTGAGAGAGCTGCGCCGCTGCGAGAG 1139
QY 1060 GCGCTGCGGACAGCTTCGATGAGAGGAGGAGTGAAGTGGCTGCTCAACCGCATCCCTGCG 1119
Db 1140 GCGCTGCGGACAGCTTCGATGAGAGGAGGAGTGAAGTGGCTGCTCAACCGCATCCCTGCG 1199
QY 1120 CGGGACATGCCCCCACTCATTCGACCAACCGCTTCACGGCCAGCTTCCAGGGCATC 1179
Db 1200 CGGGACATGCCCCCACTCATTCGACCAACCGCTTCACGGCCAGCTTCCAGGGCATC 1259
QY 1180 GTGGATCGTACGGCGTGGCGCTACAGAGAGGTCAACCCCGCTCCCTACAGCATCATC 1239
Db 1260 GTGGATCGTACGGCGTGGCGCTACAGAGAGGTCAACCCCGCTCCCTACAGCATCATC 1319
QY 1240 ACCTTCCCTTCTGTTGCTGTGATGTTTCGGGATGTGGGCAACGGGCTGCTCATGTTTC 1299
Db 1320 ACCTTCCCTTCTGTTGCTGTGATGTTTCGGGATGTGGGCAACGGGCTGCTCATGTTTC 1379
QY 1300 CTCTTGCGCCTGCGCATGCTCTTGGGAGAACCGACCGGCTGTGAAGCCCGCGCAGAAC 1359
Db 1380 CTCTTGCGCCTGCGCATGCTCTTGGGAGAACCGACCGGCTGTGAAGCCCGCGCAGAAC 1439
QY 1360 GAGATCTGGCAGACTTCTTACGGGCGGCTACCTGCTCTGCTATAGGGCTGTTCTCC 1419
Db 1440 GAGATCTGGCAGACTTCTTACGGGCGGCTACCTGCTCTGCTATAGGGCTGTTCTCC 1499
QY 1420 ATCTACACCGGCTTATCTACACAGAGTCTTTCAGTGGCGGCAACAGCATCTTCCCTCG 1479
Db 1500 ATCTACACCGGCTTATCTACACAGAGTCTTTCAGTGGCGGCAACAGCATCTTCCCTCG 1559
QY 1480 GCGTGAAGTGTGGCGCCATGCGCAACAGTGTGGCTGAAGTGAATCTTGGGCCAG 1539
Db 1560 GCGTGAAGTGTGGCGCCATGCGCAACAGTGTGGCTGAAGTGAATCTTGGGCCAG 1619
QY 1540 CACAGATGCTTACCTGATCCCAACGCTCACGGGTCTTCTGGGACCCCTACCCCTTT 1599
Db 1620 CACAGATGCTTACCTGATCCCAACGCTCACGGGTCTTCTGGGACCCCTACCCCTTT 1679
QY 1600 GGCATCGATCTTATTTGGAGCCTGGCTGCGCAACCACTTGAGCTTCTCAACTCTTCAAG 1659
Db 1680 GGCATCGATCTTATTTGGAGCCTGGCTGCGCAACCACTTGAGCTTCTCAACTCTTCAAG 1739
QY 1660 ATGAAGATGCTGCTCATCTCGGCGTGTGCAATGGCCCTTTGGGCTGCTCGAGTCT 1719
Db 1740 ATGAAGATGCTGCTCATCTCGGCGTGTGCAATGGCCCTTTGGGCTGCTCGAGTCT 1799
QY 1720 TTCAACCAAGTGCATTTGGCCAGAGGCAACCGGCTGCTGAGAGCGTTCGCGAGCTC 1779
Db 1800 TTCAACCAAGTGCATTTGGCCAGAGGCAACCGGCTGCTGAGAGCGTTCGCGAGCTC 1859
QY 1780 ACCTTCTGCTGGGACTCTTCCGTTTACCTCGTGTCTTCTAGTCTATCTACAAGTGGCTGT 1839
Db 1860 ACCTTCTGCTGGGACTCTTCCGTTTACCTCGTGTCTTCTAGTCTATCTACAAGTGGCTGT 1919
QY 1840 GTCTGGGCTGCGAGGCGCCCTCG---CCAGATCTCTCATCTCATCTTCAATCAATGTTTC 1896
Db 1920 GTCTGGGCTGCGAGGCGCCCTCGGGCCCCCAGCATCTCTCATCTTCAATCAATGTTTC 1979
QY 1897 CTCTTCTCCACAGCCCCAGCAACAGGCTGCTTACCCCGGCGAGGAGTGTCCAGGCC 1956
Db 1980 CTCTTCTCCACAGCCCCAGCAACAGGCTGCTTACCCCGGCGAGGAGTGTCCAGGCC 2039

RESULT 10
US-09-949-016-1330
; Sequence 1330, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1330
; LENGTH: 2457
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1330
```

| Query Match | | | | | | | | | |
|---|------|---|------|--|--|--|--|--|--|
| Best Local Similarity 77.7%; Score 2051; DB 3; Length 2457; | | | | | | | | | |
| Matches 2088; Conservative 0; Mismatches 0; Indels 3; Gaps 1; | | | | | | | | | |
| Qy | 496 | GGGGCCTCAGAGAGGACGCCCTCTCCAGCCCCCGGGGGCCGACACAGGACCTGAGG | 555 | | | | | | |
| Db | 327 | GGGGGCTTGTGGGGAGGCGAGAGGTTGGAGCAGCCCTGCCCCAGCCCGCTGCC | 386 | | | | | | |
| Qy | 556 | GTCACACTTTGTGGCAGGTGCCGTGGAGCCCCACAAAGGCCCTGCGCTAGAGCGCTGCTC | 615 | | | | | | |
| Db | 387 | GCCAGCTTTGTGGCAGGTGCCGTGGAGCCCCACAAAGGCCCTGCGCTAGAGCGCTGCTC | 446 | | | | | | |
| Qy | 616 | TGGAGGCGCTGCCCGCGCTTCCCTCATTTGCCAGCTTTCAGGGAGCTGGAGCAGCCGCTGGAG | 675 | | | | | | |
| Db | 447 | TGGAGGCGCTGCCCGCGCTTCCCTCATTTGCCAGCTTTCAGGGAGCTGGAGCAGCCGCTGGAG | 506 | | | | | | |
| Qy | 676 | CACCCCGTGACGGGGGAGCCAGCCACGTGGATGACCTTCTCATCTCTACTTGGGGTGAG | 735 | | | | | | |
| Db | 507 | CACCCCGTGACGGGGGAGCCAGCCACGTGGATGACCTTCTCATCTCTACTTGGGGTGAG | 566 | | | | | | |
| Qy | 736 | CAGATCGGACAGAAAGATCCGCAAGATCACGGACTGCTTCCACTGCCACGCTTCCCGCTTT | 795 | | | | | | |
| Db | 567 | CAGATCGGACAGAAAGATCCGCAAGATCACGGACTGCTTCCACTGCCACGCTTCCCGCTTT | 626 | | | | | | |
| Qy | 796 | CTGCAGCAGGAGGAGGCCGCGCTCGGGGCCCTGCAGCAGCTGCAACAGCAGAGCCAGGAG | 855 | | | | | | |
| Db | 627 | CTGCAGCAGGAGGAGGCCGCGCTCGGGGCCCTGCAGCAGCTGCAACAGCAGAGCCAGGAG | 686 | | | | | | |
| Qy | 856 | CTGCAGGAGTCTCTCGGGGAGACAGAGGGTTCTTGAGCCAGGTGCTAGGCCGGGTGCTG | 915 | | | | | | |
| Db | 687 | CTGCAGGAGTCTCTCGGGGAGACAGAGCGTTCTTGAGCCAGGTGCTAGGCCGGGTGCTG | 746 | | | | | | |
| Qy | 916 | CAGCTGCTGCCGCCAGGGCAGGTGCAGTCCACAAAGATGAAGGCCGTGTACTTGGGCCCTG | 975 | | | | | | |
| Db | 747 | CAGCTGCTGCCGCCAGGGCAGGTGCAGTCCACAAAGATGAAGGCCGTGTACTTGGGCCCTG | 806 | | | | | | |
| Qy | 976 | AACCAAGTGACGCTGAGCACCAACAGTGCCTCATTTGCCGAGGCGCTGTGCTCTGTG | 1035 | | | | | | |
| Db | 807 | AACCAAGTGACGCTGAGCACCAACAGTGCCTCATTTGCCGAGGCGCTGTGCTCTGTG | 866 | | | | | | |
| Qy | 1036 | CGAGACTGCCCGCCCTGCAGGAGGCCCTTGGGGACAGCTCGATGGAGGAGGAGTGAGT | 1095 | | | | | | |
| Db | 867 | CGAGACTGCCCGCCCTGCAGGAGGCCCTTGGGGACAGCTCGATGGAGGAGGAGTGAGT | 926 | | | | | | |
| Qy | 1096 | GCGTGCTCACCAGCATCCCTGCGGGACATGCCGCCCACTCATCTCCGACCAACCGC | 1155 | | | | | | |
| Db | 927 | GCGTGCTCACCAGCATCCCTGCGGGACATGCCGCCCACTCATCTCCGACCAACCGC | 986 | | | | | | |
| Qy | 1156 | TTACAGGCCAGCTTTCAGGGGCATCGTGGATCGCTACGGCGTGGGCCGCTACCAAGAGGTC | 1215 | | | | | | |
| Db | 987 | TTACAGGCCAGCTTTCAGGGGCATCGTGGATCGCTACGGCGTGGGCCGCTACCAAGAGGTC | 1046 | | | | | | |
| Qy | 1216 | AACCCCGCTCCCTACACCATCATCCTTCCCTTCTGTTGCTGTGATGTTTCGGGGAT | 1275 | | | | | | |
| Db | 1047 | AACCCCGCTCCCTACACCATCATCCTTCCCTTCTGTTGCTGTGATGTTTCGGGGAT | 1106 | | | | | | |
| Qy | 1276 | GTGGGCCACGGGCTGCTCATGTTCTCTTCCGCGCTGGCCATGTGCTTTCGGAGAACCGA | 1335 | | | | | | |
| Db | 1107 | GTGGGCCACGGGCTGCTCATGTTCTCTTCCGCGCTGGCCATGTGCTTTCGGAGAACCGA | 1166 | | | | | | |
| Qy | 1336 | CCGGCTGTAAAGCCCGCGCAGAACGAGATCTGGCAGACTTCTTTCAGGGGCGCTACTCTG | 1395 | | | | | | |
| Db | 1167 | CCGGCTGTAAAGCCCGCGCAGAACGAGATCTGGCAGACTTCTTTCAGGGGCGCTACTCTG | 1226 | | | | | | |
| Qy | 1396 | CTCTGCTTATGGGCGCTGTTTCTCOATCTACACCGGCTTCATCTACACGAGTGCTTCAGT | 1455 | | | | | | |
| Db | 1227 | CTCTGCTTATGGGCGCTGTTTCTCOATCTACACCGGCTTCATCTACACGAGTGCTTCAGT | 1286 | | | | | | |
| Qy | 1456 | CGGCGCACAGCATCTTCCCTCTGGGCTGAGTGTTGGCCGCGCATGGCCACCAAGTCTGGC | 1515 | | | | | | |
| Db | 1287 | CGGCGCACAGCATCTTCCCTCTGGGCTGAGTGTTGGCCGCGCATGGCCACCAAGTCTGGC | 1346 | | | | | | |

| | | | | | |
|----|------|----------------------------------|-------------------------------------|------------------------|------|
| Qy | 1516 | TGGAGTCATGCAATTCCTGGGCCAGACACA | CGATGCTTACCTCTGGATCCAAACGTCACCGGT | 1575 | |
| Db | 1347 | TGGAGTCATGCAATTCCTGGGCCAGACACA | CGATGCTTACCTCTGGATCCAAACGTCACCGGT | 1406 | |
| Qy | 1576 | GTCTTCTCTGGGACCCCTACCCCTTTGGCAT | CGATCCTATTTTGGAGCCTGGCTGCCAACCC | 1635 | |
| Db | 1407 | GTCTTCTCTGGGACCCCTACCCCTTTGGCAT | CGATCCTATTTTGGAGCCTGGCTGCCAACCC | 1466 | |
| Qy | 1636 | TTGAGCTTCTCAACTCTCTTCAAGATGAAGAT | GTCCGTCATCTCTGGGCGCTCGTGCAATG | 1695 | |
| Db | 1467 | TTGAGCTTCTCAACTCTCTTCAAGATGAAGAT | GTCCGTCATCTCTGGGCGCTCGTGCAATG | 1526 | |
| Qy | 1696 | GCCTTTGGGCTGTGCTCGAGCTTCCAAACCA | CCACGTGCATTTTGGCCAGAGGCCACCGGCTG | 1755 | |
| Db | 1527 | GCCTTTGGGCTGTGCTCGAGCTTCCAAACCA | CCACGTGCATTTTGGCCAGAGGCCACCGGCTG | 1586 | |
| Qy | 1756 | CTGCTGAGACGCTGCGGAGCTCACCTTCTCT | GCTGGGATCTCTTCCGTTACCTCGTGTTC | 1815 | |
| Db | 1587 | CTGCTGAGACGCTGCGGAGCTCACCTTCTCT | GCTGGGATCTCTTCCGTTACCTCGTGTTC | 1646 | |
| Qy | 1816 | CTAGTCATCTACAAGTGGCTGTGTCTGGGCT | GCCAGGGCCGCTCG---CCAGCATC | 1872 | |
| Db | 1647 | CTAGTCATCTACAAGTGGCTGTGTCTGGGCT | GCCAGGGCCGCTCGGCCCCAGCATC | 1706 | |
| Qy | 1873 | CTCATCACTTCATCAACATGTTCCTCTTCTCC | CACAGCCCCAGCAACAGGCTGCTCTAC | 1932 | |
| Db | 1707 | CTCATCACTTCATCAACATGTTCCTCTTCTCC | CACAGCCCCAGCAACAGGCTGCTCTAC | 1766 | |
| Qy | 1933 | CCCCGAGAGTGCTCAGGCCACGCTGGTGGT | CTTGCCCTTGGCCATGGTGCCCATC | 1992 | |
| Db | 1767 | CCCCGAGAGTGCTCAGGCCACGCTGGTGGT | CTTGCCCTTGGCCATGGTGCCCATC | 1826 | |
| Qy | 1993 | CTGCTGCTTGGCACACCCCTGCACCTGCTGC | ACCGCACCGCCGCCCTGCCGAGGAGG | 2052 | |
| Db | 1827 | CTGCTGCTTGGCACACCCCTGCACCTGCTGC | ACCGCACCGCCGCCCTGCCGAGGAGG | 1886 | |
| Qy | 2053 | CCGCTGACCGACAGGAGGAAAAA | CAAGGCCGGGTGTGTGCACTGCTGACGATCTGTG | 2112 | |
| Db | 1887 | CCGCTGACCGACAGGAGGAAAAA | CAAGGCCGGGTGTGTGCACTGCTGACGATCTGTG | 1946 | |
| Qy | 2113 | AATGGCTGGAGCTCCGATGAGGAAAAAGGCA | GGGGGCTTGATGATGAGAGGAGGCCGAG | 2172 | |
| Db | 1947 | AATGGCTGGAGCTCCGATGAGGAAAAAGGCA | GGGGGCTTGATGATGAGAGGAGGCCGAG | 2006 | |
| Qy | 2173 | CTGCTCCCTCTCGAGGTGCTCATGCA | CCAGGCCATCCACACCATCGAGTTCTGCCTGGGC | 2232 | |
| Db | 2007 | CTGCTCCCTCTCGAGGTGCTCATGCA | CCAGGCCATCCACACCATCGAGTTCTGCCTGGGC | 2066 | |
| Qy | 2233 | TGCGTCTCCAACACCGCTCTTACCTACCT | GTGCGCTGTGGGCCCTTGAGCCTGGGCCACGCCAG | 2292 | |
| Db | 2067 | TGCGTCTCCAACACCGCTCTTACCTACCT | GTGCGCTGTGGGCCCTTGAGCCTGGGCCACGCCAG | 2126 | |
| Qy | 2293 | CTGTCCGAGGTTCTGTGGGCCATGGTGAT | CCGCAATAGGCCCTGGGCCCGGAGGTG | 2352 | |
| Db | 2127 | CTGTCCGAGGTTCTGTGGGCCATGGTGAT | CCGCAATAGGCCCTGGGCCCGGAGGTG | 2186 | |
| Qy | 2353 | GGCGTGCGGCTGTGTGTGCTGCCCATCTT | TTCGCGCTTGATGACCGTGGCT | 2412 | |
| Db | 2187 | GGCGTGCGGCTGTGTGTGCTGCCCATCTT | TTCGCGCTTGATGACCGTGGCT | 2246 | |
| Qy | 2413 | ATCCTGCTGTGTATGAGGGA | CTCTCAGCCTTCTGCA | CGCCCTGGGGCTGCACTGGGTG | 2472 |
| Db | 2247 | ATCCTGCTGTGTATGAGGGA | CTCTCAGCCTTCTGCA | CGCCCTGGGGCTGCACTGGGTG | 2306 |
| Qy | 2473 | GAATTCGAGAACAAAGTTCTACTCAGGCA | CGGGCTACAGCTGAGTCCCTTCACTTTCGCT | 2532 | |
| Db | 2307 | GAATTCGAGAACAAAGTTCTACTCAGGCA | CGGGCTACAGCTGAGTCCCTTCACTTTCGCT | 2366 | |
| Qy | 2533 | GCACAGATCACTAGGGCCCACTGCA | GGTCTCTCCAGACCTCTTCTG | TCAACCTCTGAGGC | 2592 |
| Db | 2367 | GCCACAGATCACTAGGGCCCACTGCA | GGTCTCTCTCCAGACCTCTTCTG | TCAACCTCTGAGGC | 2426 |
| Qy | 2593 | AGGAGAGGAAATAAAGACGGTCCG | CCCTGGCA | 2623 | |

| | | |
|----|------|---|
| Db | | 2427 AGGAGAGGAATAAAGACGGTCCGCCCTGGCA 2457 RESULT 11 US-09-949-016-1331 ; Sequence 1331, Application US/09949016 ; Patent No. 6812339 ; GENERAL INFORMATION: ; APPLICANT: VENTER, J., Craig et al. ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF ; FILE REFERENCE: CL001307 ; CURRENT APPLICATION NUMBER: US/09/949,016 ; PRIOR FILING DATE: 2000-04-14 ; PRIORITY APPLICATION NUMBER: 60/241,755 ; PRIOR FILING DATE: 2000-10-20 ; PRIOR APPLICATION NUMBER: 60/237,768 ; PRIOR FILING DATE: 2000-10-03 ; PRIOR APPLICATION NUMBER: 60/231,498 ; PRIOR FILING DATE: 2000-09-08 ; NUMBER OF SEQ ID NOS: 207012 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 1331 ; LENGTH: 2457 ; TYPE: DNA ; ORGANISM: Human US-09-949-016-1331 |
| | | Query Match 77.7%; Score 2051; DB 3; Length 2457; Best Local Similarity 98.0%; Pred. No. 0; Matches 2088; Conservative 0; Mismatches 40; Indels 3; Gaps 1; |
| QY | 496 | GSGGCGCTCAGAGAGGACGCCCTGTCTCCAGGCCCCCGGGGGCGGCACCAGGACTTGAGG 555 |
| DB | 327 | GGGGGCGCTGGTGGGGAGGACAGGGCAGAGGTTGGAGCAGCGCTTGCCCGAGCCCCGTGGCC 386 |
| QY | 556 | GTCAAATTTTGTGGCAGGTGCCGTGGAGGCCCAAGAAGGCCCTTGCCCTAGAGCGGCTGCTC 615 |
| DB | 387 | GCCAGCTTTTGTGGCAGGTGCCGTGGAGGCCCAAGAAGGCCCTTGCCCTAGAGCGGCTGCTC 446 |
| QY | 616 | TGAGGGCGCTGCGCGGCTTCCTCATATGCCAGCTTTCAGGGAGCTGGAGCACGCCGCTGGAG 675 |
| DB | 447 | TGAGGGCGCTGCGCGGCTTCCTCATATGCCAGCTTTCAGGGAGCTGGAGCACGCCGCTGGAG 506 |
| QY | 676 | CACCCCGTCAGCGGGGAGCCAGCCAGCTGGATGACCTTCCATCTCCTACTGGGGTGAG 735 |
| DB | 507 | CACCCCGTCAGCGGGGAGCCAGCCAGCTGGATGACCTTCCATCTCCTACTGGGGTGAG 566 |
| QY | 736 | CAGATCGGACAGAAATCCGCAAGATCAGGAGCTGCTTCCAATGCGCACGCTCTTCCCCTTT 795 |
| DB | 567 | CAGATCGGACAGAGATCCGCAAGATCAGGAGCTGCTTCCAATGCGCACGCTCTTCCCCTTT 626 |
| QY | 796 | CTGACGACGAGAGAGGCCCGCTTCCGGGCGCTTGACGAGCTGCAACAGCAGACGAGCAGGAG 855 |
| DB | 627 | CTGACGACGAGAGAGGCCCGCTTCCGGGCGCTTGACGAGCTGCAACAGCAGACGAGCAGGAG 686 |
| QY | 856 | CTGACGAGGCTCTCGGGGAGACAGAGCGGTTCTTGAGCCAGGCTTAGGCCGGGTGCTG 915 |
| DB | 687 | CTGACGAGGCTCTCGGGGAGACAGAGCGGTTCTTGAGCCAGGCTTAGGCCGGGTGCTG 746 |
| QY | 916 | CAGCTGCTCCGCGCAGGGCAGGTGAGGTGCCAAGATGAAGGCGCGTGTACTTGGCCCTG 975 |
| DB | 747 | CAGCTGCTCCGCGCAGGGCAGGTGAGGTGCCAAGATGAAGGCGCGTGTACTTGGCCCTG 806 |
| QY | 976 | AACAGGTGAGCGTGAGCACCAACCAAGTGCCTCATTTGCCGAGGCGCTGGTCTCTGTG 1035 |
| DB | 807 | AACAGGTGAGCGTGAGCACCAACCAAGTGCCTCATTTGCCGAGGCGCTGGTCTCTGTG 866 |
| QY | 1036 | CGAGACTTCCCGCCCTGACGAGGCGCTTGCGGACAGCTCGATGGAGGAGGAGTGAGT 1095 |
| DB | 867 | CGAGACTTCCCGCCCTGACGAGGCGCTTGCGGACAGCTCGATGGAGGAGGAGTGAGT 926 |

Db 2007 CTGCTCCCTCCGAGGTGCTATGACACAGGCCATCCACCATCGAGTTCTGCTTGGC 2066
Qy 2233 TGGGTCTCCAAACACCGCTCTCTACTGCGCTGTGGCCCTTGAGCCTGGGCCACAGCCAG 2292
Db 2067 TGGGTCTCCAAACACCGCTCTCTACTGCGCTGTGGCCCTTGAGCCTGGGCCACAGCCAG 2126
Qy 2293 CTGTCGAGGTCTGTGGCCATGCTGATGGCATAGCCTGGGCCCTGGGCCGGGAGTG 2352
Db 2127 CTGTCGAGGTCTGTGGCCATGCTGATGGCATAGCCTGGGCCCTGGGCCGGGAGTG 2186
Qy 2353 GGGTGGCGCTGTGTGCTGCTGCCCATCTTTGGCCGCTTTGGCGTGATGACCGTGGCT 2412
Db 2187 GGGTGGCGCTGTGTGCTGCTGCCCATCTTTGGCCGCTTTGGCGTGATGACCGTGGCT 2246
Qy 2413 ATCTGTGTGTGTGAGGAGTCTCAGCTTCTCAGCTTCTGCAACCGCTTGGCGCTGCACTGGGTG 2472
Db 2247 ATCTGTGTGTGTGAGGAGTCTCAGCTTCTCAGCTTCTGCAACCGCTTGGCGCTGCACTGGGTG 2306
Qy 2473 GAATTCAGAACAAAGTTCTACTCAGGACCGGCTCAGAGCTGAGTCCCTTCACTTGGCT 2532
Db 2307 GAATTCAGAACAAAGTTCTACTCAGGACCGGCTCAGAGCTGAGTCCCTTCACTTGGCT 2366
Qy 2533 GCCACAGATCAGTAGGCCCACTGCACTGCTGCAAGCTTCTTCTGACCTCTGAGGC 2592
Db 2367 GCCACAGATCAGTAGGCCCACTGCACTGCTGCAAGCTTCTTCTGACCTCTGAGGC 2426
Qy 2593 AGGAGAGGAATAAGACGGTCCGCCCTGGCA 2623
Db 2427 AGGAGAGGAATAAGACGGTCCGCCCTGGCA 2457

RESULT 12
5196526-2
; Patent No. 5196526
; APPLICANT: BEAMAN, KENNETH D.
; TITLE OF INVENTION: CDNA CLONE FOR T-CELL SUPPRESSOR INDUCER
; FACTOR
; NUMBER OF SEQUENCES: 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/442,116
; FILING DATE: 28-NOV-1989
; SEQ ID NO:2:
; LENGTH: 2945
5196526-2

Query Match 28.1%; Score 741.2; DB 10; Length 2945;
Best Local Similarity 58.5%; Pred. No. 1.7e-125;
Matches: 1486; Conservative 0; Mismatches 973; Indels 81; Gaps 8;

Qy 66 CCATGGGCTCCATGTTCCGAGGAGGAGGAGTGGCCCTGTGCTCCAGCTTTTCTGCCCAAG 115
Db 17 CCATGGGCTCTCTCTTCCGAGGAGAGAGTGTGCTTGGCGAGCTTCTTCTGCACTGCG 76
Qy 116 CGGCTGCCCTACACTGGGTGAGTGGCTGGCGAGCTGGCCCTGTGGATTGAGAGACC 175
Db 77 GCACGGCTACGAGTGTGTAGGCGCTGGCGAGAGGAGGCTGGTGCAGTTCGAGACC 136
Qy 176 TCAACGCTCGGTGAGCGCTTCCAGAGACGCTTTGTGTGTGATGTTGGCGCTGTGAGG 235
Db 137 TCATCAAAATGTAAGTCTTTTCAAGAAATTCGTCGTGAGGTAAGAGAGTGTGAAG 196
Qy 236 AGCTGAGAGACCTTCACTTCTGAGAGAGGAGTGGCGGCTGGGCTGGTCTGCTGC 295
Db 197 AGCTGAGAGACCTTCACTTCTGAGAGAGGAGTGGCGGCTGGGCTGGTCTGCTGC 256
Qy 296 CCCCAGCAAGAGGAGGCTGCGGCAACCCAGGAGCTGCTGCGGCTGAGGAGG 355
Db 257 CTGAAGAGAGGAGGAGTCTTCCGCAACCTCTTAAACATGTCTTAAACATGTCTTAAAC 316
Qy 356 AGACGAGAGGCTGCGGAGGCTGCGGAGTGTGCGGAGCAACAGGAGGCTGCGG 415
Db 317 AGCTGAGAGGCTGCGGAGGCTGAGAGTGTGAGAGTGTGAGAGAGGAGTGTGAGAG 376

Qy 416 CCAGGTGACACAGCTGACGCTCCAGCCGCCGCTGTC-----TACGCC 457
Db 377 AGAACCTGTGTGAGCTGTGTGAGTACACCCACATGCTGAGGGTACGAAAGACCTTCTCTCA 436
Qy 458 AGGGCCATGAACCTCAGCTGGCAGCCGCCACACAGATGGGGCTCAGAGAGAGAGCCCC 517
Db 437 AGCGAATGTGAGATTTGAGCCCACTGATGAAGATTTCCCGCTTATAGAAACGACTTCTC 496
Qy 518 TGCTCCAGGCCCCCGGGGGCCGACACAGGACCTGAGGGTCAACTTTGTGGCAGGTGCG 577
Db 497 TGTGGACTACAGCTGATGACGCGCTGGCGCAAGCTGGGATTCGTTCTGCGCTGA 556
Qy 578 TGGAGCCCAAGGCCCTTGCCTTGTAGAGCGCTGCTCTGTGAGAGGCTTGCCTGGCTTCC 637
Db 557 TTAGCAGAGGAGGAGTGAAGCATTTGAAAGATGCTGTGGCGGGCTGCAAGAGGTACA 616
Qy 638 TCATTCGACGCTTTCAGGAGCTGAGCAGCCGCTGGAGCACCCGCTGACGGCGAGCCAG 697
Db 617 CCATTTGACCTATGACAGCTGAGCGAGTGCCTGGAGGACCTTGAACCGGGTGAAGTCA 676
Qy 698 CCACGTGGATGACCTTCTCTCTCTCTGAGGCTGAGCAGATCGGACAGAAATCGGCA 757
Db 677 TAAAGTGTACGTGTTCTCTGATTTCTTTTGGGAGAGCAGATTTGGCCACAAGTTAAGA 736
Qy 758 AGATCAGGAGTCTCTCAGTGCACGCTTCTTCCGCTTCTGTGACAGAGAGGAGCCGCC 817
Db 737 AGATATGTGACTGTTTACCACTGCGCACATCTACCGTATCCAAACACCGCTGAGAGCGCA 796
Qy 818 TCGGGGCCCTGACAGAGCTGCAACAGCAGAGCCAGGAGCTGACAGGAGTCTCTGGGAGA 877
Db 797 GGGAGATTCAGGAGGGGCTCAATACTCGGATCAAGACCTTTACACGCTGCTGCACAAA 856
Qy 878 CAGAGCGGTTCTGAGCCAGGTGCTAGGCGGGTGTGTCAGCTGCTGCCCGCAGGCGAG 937
Db 857 CGAGGACTATCTGCGCAAGTCTCTGCAAAAGCTGCCGAGTCCGTGTGAGCCGTGG 916
Qy 938 TGCAGTTCACAAGATGAAGGCGCTGTACTGCTGCCCTTGAAACAGTGCAGGTGAGACCA 997
Db 917 TCCAGGTGAGGAGATGAAGGCCATCTACCACTGTCTCAACATGTGAGCTTTGACGTCA 976
Qy 998 CGCACAGTGCCTCATTTCCGAGGCGCTGTGCTGTCGAGAGCTTCCCGCTGAGG 1057
Db 977 CCAACAGTGCCTCATAGCTGAGGTCTGTGTGCGGAGGTGAGCTTCCAGGCTTGGCA 1036
Qy 1058 AGGCGCTGCGGAGCAGCTCGATGAGGAGGGAG-----TGAGTGCCTGGCTCACCGCA 1111
Db 1037 GAGCACTGGAGGAGGCTCGAGAGAGAGCGGAGCTACATCCCTCGTTTATGAAACAA 1096
Qy 1112 TCCCTTGGCGGAGCATGCCCCCACCACATCATCCGACCAACCGCTTCAAGCCAGCTTCC 1171
Db 1097 TCCCTACGAAAGAAACACCCCCACTCTGATCCGACCAACAAATTCACCTGAAGGTTTCC 1156
Qy 1172 AGGGCATCTGGATCCCTAGCGGTGGCGCTTACAGGAGGTCAACCCCGCTCCCTACA 1231
Db 1157 AGAACATCTGGATGCTTATGGAGTCCGGAGCTACAGAGAGTGAACCCAGCTCTCTTTA 1216
Qy 1232 CCATCATCACTTCCCTTCTCTGCTGTGATGTTGGGGATGTGGGCAACCGGCTGC 1291
Db 1217 CCATCATCACTTCCCTTCTCTGCTGTGATGTTGGGCACTTTGGGCAAGGCTTTG 1276
Qy 1292 TCATGTTCTTCTTGGCCCTGGCCATGCTCTTGGGAGAACCGACCGGCTGTGAAAGCCG 1351
Db 1277 TCATGTTCTGTTTGGCTTCTTACTGTTGTTAAATGAGAAATCAACCCAGACTAAGCCAGT 1336
Qy 1352 CGCAGAACAGATCTGGCAGACTTCTTCAAGGGCGCTTACCTGCTCTGCTTATGGGCC 1411
Db 1337 CACAG---GAGATCTCTTAGGATGTTCTTTTGAACCGGCTATATCTGCTGCTGATGGGG 1393
Qy 1412 TGTTCCTCACTACACCGGCTTCACTACAAAGAGTGTCTTCACTGCGGCCACAGCATCT 1471
Db 1394 TGTTCCTGTGTACCGGCTCATCTACAAAGCTGCAAGAGGTTCTCTGTGAACCTCT 1453

QY 820 GGGGCCCTGCGAGCTGCAACAGCAGAGCCAGAGCTGCGAGAGGTCCTTCGGGGAGACA 879
DB 781 GAGATTTCAGGAGGGGCTCAATACCTCGGATCCAGACCTTTACACGGTGTGCAAAAACG 840
QY 880 GAGCGGTTCTGAGCCAGGTGCTAGGCGGGTGTGCTGAGCTGCTGCGCCAGGCGAGTG 939
DB 841 GAGACTATCTGCGGCAAGTGCTCTGCAAGCTCCGAGTCCGCTGCTGAGCGCGTGTGTC 900
QY 940 CAGGTCCACAAGATGAAGCGCGTGTACTTGGCCCTGAACCAAGTGCAGCTGAGCACCACG 999
DB 901 CAGGTGAGGAAGATGAAGGCCATCTACCAATGCTCAACATGTGTCAGCTTTGACGTCAAC 960
QY 1000 CACAGTGCCTCATTTGCGAGGCTGTGCTGTGCTGAGACCTTGGCCGCGCTGCGAGG 1059
DB 961 AACAGTGCCTCATAGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1060 GCCCTGCGGACAGCTCGATGGAGGAGGAG-----TGAGTCCGCTGGCTCACCGCATC 1113
DB 1021 GCATGAGGAGAGGCTCGAGAGAGCGGAGCTCAATCCCTCGTTCAATGAACAACATC 1080
QY 1114 CCTGCGGAGCATGCCCGCCACACTCATCGCAACCAACCGCTTTCAGGCGCAGCTTCCAG 1173
DB 1081 CCTACGAAAGAAACACCCCGCCACTCTGATCGCACCAACAAATTCATCTGAAGGTTTCCAG 1140
QY 1174 GGCATGCTGATGCTAGCGGCTGGCGGCTACAGGAGTCAACCCCGCTCCCTACAC 1233
DB 1141 AACATGCTGATGCTATGAGGCTGGGAGCTACAGAGAAGTGAACCCAGCTCTCTTTACC 1200
QY 1234 ATCATACCTTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1293
DB 1201 ATCATACCTTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1294 ATGTTCTCTTCCGCTTGGCGCATGCTTTCGGAGAACCGACCGGCTGTGAAGCGCG 1353
DB 1261 ATGTTCTGTTTGGCTTCTTACTGGTGTAAATGAGATCAACCCAGACTAAGCCAGTCA 1320
QY 1354 CAGAACGAGATGCGGAGACTTTCTTCAGGGGCGCTACTGCTGCTGCTGCTGCTGCTGCT 1413
DB 1321 CAG---GAGATCCTTAGGATGTTCTTTGACGGCGGCTATATCTGCTGCTGCTGCTGCTG 1377
QY 1414 TTCTCCATACACCGGCTTCACTCAACAGAGTCTTCACTGCGGCGCACCAGCATCTTC 1473
DB 1378 TTCTGCTGCTACACCGGCTTCACTCAACAGACTTCTTCTCAAGTCTGGAACCTCTTT 1437
QY 1474 CCCTCGGGCTGAGTGTGGCGGCTACCGCAACCGAGTCTGGC----- 1515
DB 1438 GGTCTCGGGTGAACGATATGTGCAATGTACAGCTCCAGCCACTCTCCAGAGGAGCAGAG 1497
QY 1516 -----TGGAGTATGATTCCTGGCCGACGACAGATGCTTACCTGGATCCC 1563
DB 1498 AAGATGGTACTTTTGAATGACAGCACCATCAGGCACAGCAAACTTTGCAAGCTGGACCCG 1557
QY 1564 AAGCTCACCGGCTCTTCTGCGGACCTTACCCCTTTGGCATCGATCTTATTTGAGGCTG 1623
DB 1558 AATATCCCTGGAGTTTTCGAGGCGGCTTACCTTTTGGCATGATCTATTTGGAACCTG 1617
QY 1624 GCTGCCAACCACTTGAAGCTTCTCAACTCTTCAAGATGAAGTGTCCGTCATCTCTGGG 1683
DB 1618 GCCACAAACCGCTCAGCTTCTCACTCTTCAAGATGAATGCTGCTGATTTAGGA 1677
QY 1684 GTGCTGACATGCGCTTTTGGGGTGTCTCGGAGTCTTCAACCAAGTGAAGTCTTTGGCCAG 1743
DB 1678 ATTTTTCATGACTTTTGGTGTGTTCTGGGAATATTTAACCACTTGCACCTTTAGGA 1737
QY 1744 AGGCACCGGCTGCTGCTGAGAGCGCTCGGAGCTCACCTTCTGCTGGGACTCTTCGCT 1803
DB 1738 AAGTTCAACGCTTACCTGGTCTCGGTCCTGAGATCTGTTTCAATGCTCTGCTGCTTGGG 1797
QY 1804 TACTGCTGTTCTAGTCACTACAGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
DB 1798 TACTGATTTTCATGATCATTTACAGTGGCTGGCATACTTCGGAGAGACTTCGAGAGAG 1857

QY 1861 TCGCCAGCATCTCTCATCTTCACTTCAACATGTTCTTCTTCTCCCAAGCCCCAGCAAC 1920
DB 1858 GCCCGAGCATCTCTGATCGAGTTTCAATTAACATGTTCTTCTTCCCAACAGCAACACAC 1917
QY 1921 AGGCTGCTTACCCCGGAGGAGTGTCCAGSCACGCTGTGTGTCTTCTGCGCTTGGCC 1980
DB 1918 GGG---CTTACCCAGGGCAGGCACAGTCCAGAGATGTTGTGGTCTTCAACGGTGTG 1974
QY 1981 ATGCTGCCATCTCTGCTTGGCACACCTTGCACCTGTGTGCAACCGCCACCGCGCGCC 2040
DB 1975 GCTGTCCCGTCTTCTTCTTAGGAAAGCGCTTTTCTGTGTGTGCTGCAACCGGCGC 2034
QY 2041 CTGGG-----GAGGAGCCCGCTGACCGACAGAGGAAACAAGCCGGGTGTGCTG 2091
DB 2035 AATTGCTTTTGGCATGCGCGAGCGGTTACACACTCGTGAAGAGGACAGCAGGAGAGAG 2094
QY 2092 GACCTGCTGACGATCTGTGAATGGCTGAGCTCCGATGAGGAAGGAGGAGGCGCTG 2151
DB 2095 GTGTCTCTTCTGGGCAACAGGACATAGAAGGGGCAACAGCCGATGGAAGAGGCTGC 2154
QY 2152 GATGATGAAGAGGAGCGGAGCTCGTCCCTCCGAGGTGCTCATGCAACAGGCCATCCAC 2211
DB 2155 GGAGAGTGAAGTGTGAGGAGTTTAACTTCGGGAGATCTGTATGACGCGAGGCCATCCAC 2214
QY 2212 ACCATGAGTTCGCTGCGGCTGCTCTCAACACCGCTCTTACCTGCGCTGTGGGCG 2271
DB 2215 TCCATTTGAGTACTGCTCGCTCGGCTGCTCAACACCGCTCTTACCTGAGGCTCTGGGCA 2274
QY 2272 CTGAGCTGCGCCACGCGCAGCTGTCGAGGTTCTGTGGGCGCATGTGATGCGCATAGGC 2331
DB 2275 CTGAGCTGCGCCATGACAGCTCTGTGATGCTGTGGGCGCATGTGATGCGCGTGGG 2334
QY 2332 CTGGGCTGCGGCGGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2391
DB 2335 CTG-----CGTGTGGACACCACTATGCGGCTTCTGTGCTGCTGCTGCTGCTGCT 2385
QY 2392 TTTGCGGTGATGACCGTGGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2451
DB 2386 TTTGCGATTTTGAACCAATTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2445
QY 2452 GCCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2511
DB 2446 GCATAGGCTTCACTGGGTAGAAATTTCAACAATAATTTACGTTGTGCGAGGCCACAG 2505
QY 2512 CTGAGTCCCTTCACTTTC 2529
DB 2506 TTTGTTCCCTTCTCTTC 2523

RESULT 14

US-09-799-451-895
; Sequence 895, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungling
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and

;/ TITLE OF INVENTION: Polypeptides
;/ FILE REFERENCE: 803
;/ CURRENT APPLICATION NUMBER: US/09/799,451
;/ CURRENT FILING DATE: 2001-03-05
;/ NUMBER OF SEQ ID NOS: 948
;/ SOFTWARE: pt_FL_genes Version 2.0
;/ SEQ ID NO 895
;/ LENGTH: 3062
;/ TYPE: DNA
;/ ORGANISM: Homo sapiens
;/ FEATURE:
;/ NAME/KEY: CDS
;/ LOCATION: (196)...(2715)
US-09-799-451-895

Query Match 23.4%; Score 618.6; DB 3; Length 3062;

Best Local Similarity 55.3%; Pred. No. 3.1e-103;

Matches 1410; Conservative 0; Mismatches 1099; Indels 42; Gaps 9;

```
QY      8  CGCGGACGGGACGACGAGCGGAGCGGCGGCGCAGCACACCCGGGACCATGGGCTCCA 67
DB      146  CCTGTACCCAGGACCCGCTTGAGGAAGCTGGGACCGGAGCTGGGCAAGATGGGCTCG 205

QY      68  TGTTCGGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCCCACAGCGGCTGCGCTACA 127
DB      206  TGTTCGAAGCGAGGAGATGTGTTGTCACTGTTCTCCAGGTGGAAGCTGCATATT 265

QY      128  CTTGCGTGTAGTGGCTGGGAGCTGGGCTCTGTGAGTTCAGAGACCTCAACGCTCGG 187
DB      266  GCTGTGTGCTGAGCTCGGAGAGCTCGGATTGGTTTCAGTTCAAAGATTTAAATATGAATG 325

QY      188  TGAGCGCCTTCCAGAGACCTTTGTGTTGATGTTTGGCGCTGTGAGGAGCTGGAGAAGA 247
DB      326  TGAACAGCTTTCAGAGAAATTTGTGAATGAAGTCAGAAAGGTGTGAATCACTGGAGAGAA 385

QY      248  CCTTACCTTCTTCGAGGAGGAGGTGGCGGGGCTGGGCTGGTGTCTGCCCGGCCAAAGG 307
DB      386  TCTTCGTTTCTGGAAGACGAGAT--GCAAAATGAGATTGTAGTTCAAGTTCGAGA 442

QY      308  GGAGGTGCGGACACCCACCCCGGAGCTGTGTCGGCATCCAGAGGAGAGAGCGAGCGCC 367
DB      443  AAGGCCACTGACCCCGCTCCACCGGAAATGATTACCTCGGAGACTGTTCTAGAAAAAC 502

QY      368  TGGCCCGAGGAGCTGGCGGATGTGGGGCAACACAGAGGCCCTCGCGGGCCAGCTGCACC 427
DB      503  TGAAGGAGAGTTACAGGAGGCCAACACAGAACCCAGCGCCCTTGAAACAAAGCTTCTAG 562

QY      428  AGCTGAGCTCCACGCGCGCTGTACTACGCCAGGGCCATGAACCTCAGCTGGCAGCGGCC 487
DB      563  AACTGACAGAACTGAAATACCTCTCGAAGAAACCCAGACTTCTTTGAGACGGAACCA 622

QY      488  ACACAGATGGGCGCT-----CAGAGAGAGCGCCCTGTCCAGGCCCGCGGGGGCCGC 541
DB      623  ATTTAGCTGATGATTTCTTTACTAGGAGACACTTCTGGCCCTCTCGGAGTTGAAAGCAGTGC 682

QY      542  ACCAGGACCTGAGGGTCAACTTTGTGGCAGGTGGCGTGAGGCCCAAGAGC-----CC 595
DB      683  CTGCATATATGACCGGAATTTGGGGTTTATAGCCGGTGTGATCAACAGGGAGAGATGG 742

QY      596  CTGCGCTAGAGCGCTGTCTGTGAGGGCTTGGCGGGCTTCTCTCAATGCGAGCTTCAGGG 655
DB      743  CTTCTCTTGGCGGTTACTGTGGCGAATCTGCGGAGAAACGTTGTTGAAAGTTTCAGTG 802

QY      656  AGCTGGAGCAGCGCTGGAGACCCCGTGACGGGCGAGCCAGCCAGCTGGATGACCTTCC 715
DB      803  AGATGGACCGCCCTCTGGAGGATCTCTGTGACGAAAGAAATAATTCAGAAAGACATATTCA 862

QY      716  TCATCTCTACTGGGGTGAGCAGATCGGACAGAAATCCGCAAGATCAAGGACTGCTTCC 775
DB      863  TCATATTTTACCAGGAGAGCAGCTCAGGAGAAATCAAGAGATCTGTGATGGGTTTC 922

QY      776  ACTGCCACGTCTTCCGTTTTCAGCAGCAGGAGGCGCCGCTTCGGGGCCCTTCGACAGC 835
```

```
DB      923  GAGCCACTGTCTACCCTTGCCAGAGCCTCGGGTGGAGCGCAGAGATGTTGGAGCG 982
QY      836  TGAACAGCAGAGCCAGAGCTGACAGAGTCTCGGGGAGACAGAGCGGTTCTTGAGCC 895
DB      983  TCAATGTGAGGCTGGAAGATTTAATCAACGCTCATAAACAAACAGAGTCTCACCGCCAGC 1042
QY      896  AGGTGCTAGGCGGGTGTGTCAGCTGCTGCCGCGCAGGCGAGGTGAGGTCCACAAGATGA 955
DB      1043  GCCTGCTGAGGAAGCCGCTGCCAACTCTCTGGCTCATCAAGGTGCAAGATGA 1102
QY      956  AGGCGGTGTACCTGGCCCTTGAACACAGTGCAGCGTGCAGCACACGACACCAAGTCCCTATTG 1015
DB      1103  AAGCTGTCTACCACTCTGAACTGTGCAACATCGACGTCAACCCAGCAGTGTGTCATCG 1162
QY      1016  CCGAGGCTGTGTCTGTGCGAGACCTGCGCCGCTGAGGAGGCGCTGCGGGACAGCT 1075
DB      1163  CGAGATCTGTGTTCCGGTGGCAGATGCCACACGTATCAAGAGGCACTGGAGCAAGG- 1221
QY      1076  CGATGGAGGAGGAGTGTGCGGTGCTCACCGCATCCCTGCGCGGA-----CA 1126
DB      1222  --ATGGAATAAGTGGCTCTCCATGGCCCCCATCATGACACAGTGCATCTAAACAG 1279
QY      1127  TGCCCCCCACACTCATCGGCACCAACCGCTTCACGCGCAGCTTTCAGGGCAGTCTGEGATC 1186
DB      1280  CCCCTCCACATTTAAACAGGACCAATAAATTCACAGCTGGCTTCCAGAAATATTGTTGATG 1339
QY      1187  GCTACGGGCTGGGCGCTTACAGAGAGGTCAACCCGCTCCCTTACACCAATCATCACTTCC 1246
DB      1340  CCTATGSGTTCGCGAGCTACCGGGAGATAAACCCAGCGCCCTTACACCATCATCACTTCC 1399
QY      1247  CTTCTCTGTTGCTGTGATGTTGCGGGATGTGGGCGACGGGCTGCTCATGTTCTCTTCG 1306
DB      1400  CTTCTCTGTTGCTGTGATGTTTGGAGACTGTGGTTCATGGAACCGTGTGCTCTGGCTG 1459
QY      1307  CCTCGGCATGCTCTTCGCGAGAAACCGACCGGCTGTGAAAGCCGCGCAGAACGAGATCT 1366
DB      1460  CACTTTGGATGATTTCTGAATGAGAGACGCTGCTCTCCAGAGACAGACAAATGAGATTT 1519
QY      1367  GGCAGACTTTCTTAGGGGCGCTACTGCTCTGCTCTTATGGGCTGTGTTCTCATCTACA 1426
DB      1520  GGAACACCTTCTTCCACGGGCGCTATCTGATCTCTTATGGGCACTTCTCCATCTACA 1579
QY      1427  CCGGCTTCTATCTACAAACGAGTGTCTTCACTCGCGCCACACAGCATCTTCCCTCGGCTGA 1486
DB      1580  CGGGTTTGTATCTAATGATGCTGCTTCTCAAGTCTCTTGAACATCTTTGGCTCTCTTGA 1639
QY      1487  GTGTGGCCGCTATGGCCAAACCAAGTGTGGCTGAGTGTGATTCCTGCGCCACGACACA 1546
DB      1640  GTGTCCAAACCCATGTTTCAGAAACGGCACATGGAATACTCATGTAATGGAGGAAAGTCTAT 1699
QY      1547  TGCTTACCTTGGATCCCAACGTCACCGGTGTCTTCTCTGGGA---CCCTACCCCTTTGGCA 1603
DB      1700  ATCTGACGTGGACCCAGCCATACAGGAGTGTATTTTGGAAATCCATACCCGTTTGGGA 1759
QY      1604  TCGATCTTATTTGGAGCTGGCTGCCAACCACTTCAGGCTTCTCAACTCTCTTCAAGATCA 1663
DB      1760  TTGATCCGATTTGGAACCTTGGCTTCAACAAACTCACAATTTCTGAATCTGTATAAATGA 1819
QY      1664  AGATGTCGTTCATCTGCGGCGTGTGTCACATGGCCCTTTGGGGTGTGCTCGAGTCTTCA 1723
DB      1820  AGATGTCGTGTATCTCGGAAATTTGCCAGATGGTTTGTGTCATCTCTCAGCCCTTTTCA 1879
QY      1724  ACCAGTGCATTTTGGCCAGAGGACCGGCTGCTCTGAGAGCGCTGCGGGAGCTCACCT 1783
DB      1880  ATCATATATACTTCAGAAAGAACTCTCAACATCATTTCTGCAATTTTATCTCTGAGATGATTT 1939
QY      1784  TCTCTCTGGGACTTTCGTTACCTGTTTCTAGTGTCTTCAAGTGGCTGTG---TG 1840
DB      1940  TTATCTGTGCTGTGTTGATACCTGGTTTTCATGATCATTTTCAATGGTCTGCTTGT 1999
QY      1841  TCTGGGCTGCCAGGGCGCTCGCCAGATCTCTCATCACTTCATCAACATGTTCTCTCT 1900
DB      2000  ACGTCCAGTATCTCTCAGCAGCGCCCCCAGCATCTCTCATCTCATCAACATGTTTCTGT 2059
```


Db 1235 CAGACAAACAGACTCCCCCAACTATATAACAAACCAACAGTTTACCTATGGCTTTCAG 1294
Qy 1174 GGCATCGTGATCGCTACGGCGTGGCCCGCTACAGGAGGTCAACCCCGCTCCCTACACC 1233
Db 1295 AACATAGTAGATGCTTTATGGAATTTGGAATTTACCGAGAGATAAATCCAGCTCCGTATACT 1354
Qy 1234 ATCATCACCTTCCCTTCTGTTCTGTGTGATGTTTCGGGGATGTGGGCCACCGGCTGCTC 1293
Db 1355 ATTATCACGTTTCCCTTTCTATTTGCTGTGTGATGTTTGGAGACTTCGCTCATGGCATTTTA 1414
Qy 1294 ATGTTCTCTTCGCGCTCGCCATGTTCTTGGGGAAACCGACCGGCTGTGAAAGCCGG 1353
Db 1415 ATGACCCCTTTTGTGTGTGATGTACTAGGAGAGCCGGATCCTTTCCAGAGAT 1474
Qy 1354 CAGAACGAGATCTGGCAGACTTCTTTCAGGGGCGCTACTGCTCCTGCTTATGGGCCCTG 1413
Db 1475 GAGAAATGAGATGTTTAGCACTGTGTTTCACTGGTTCGATACATTTATTTATTTGATGGGTG 1534
Qy 1414 TTCTCCATCTACACGGCTTCTATACAAACGATGCTTCAGTCGCGGCCACAGCATCTTC 1473
Db 1535 TTCTCCATGTACCTGGGCTCATCTACAATGATGCTTTTCCAGTCTCTTAATATCTTT 1594
Qy 1474 CCTCGGGCTGAGTGTGGCCCACTGGCCAAACCAAGTCTGGCTGGAGTGATGCATTTCTG 1533
Db 1595 GGGTCATCTCGAGTGACGGCCGATGTTTA--CTTATAATTGGACTGAAGAGCGCTT 1651
Qy 1534 GCCAGCACACGATGTTACCTTACCTGGATCCCAAGTCAACGGTGTCTTCTGGGACCCCTAC 1593
Db 1652 CGGGGAAACCTGTTCTACAGCTGAACCCAGCCCTCCCTGGAGTGTGTTGGTGGACCATAC 1711
Qy 1594 CCTTTGGCATCGATCTATTTTGGAGCTGGCTGCCAACCACTTGAGCTTCTCAACTCC 1653
Db 1712 CCTTTGGATTTGATCCAAATTTGGAACATTTGCTACCAATAAACTGAGTCTTGAACCTC 1771
Qy 1654 TTCAAGATGAAGATGTCCGTCTATCTCTGGCGCTCGTGCAATGCGCCTTTGGGGTGGTCTTC 1713
Db 1772 TTTAAGATGAAGATGCTGTATCTCTGTATCATCCATATGCTGTTTGGAGTCAGCCTG 1831
Qy 1714 GGAGTCTTCAACGATGCTGACTTTGGCCAGAGGACCGCGCTGCTGTGGAGACGCTGCCG 1773
Db 1832 AGTCTGTTTCAACCATATCTATTTCAAGAACCCCTGAAATATCTACTTTGGATTTATTCCT 1891
Qy 1774 GAGCTCACCTTCTGCTGGAGCTCTTCGGTATACCTCGTGTCTAGTCACTACAAGTGG 1833
Db 1892 GAAATAATCTTCATGACCTTTGTTGGCTATTTGGTTATCTTATTTTTCAGAGTG 1951
Qy 1834 CTG---TGTTGCTGGGCTGCCAGGGCCGCTCGCCAGCATCCTCATCCACTTCATCAAC 1890
Db 1952 ACGGCTATGATGCTCATACCTCTGAGAAATGCACCAAGCCTTCTGATCCATTTTATAAAC 2011
Qy 1891 ATGTTCTCTTCTCCAC---AGCCCCAGCAACAGGCTGCTACCCCGGAGGAGTG 1947
Db 2012 ATGTTCTCTTTTCTTACCCAGAGTCTGGTTATTTCAATGTTGTAATTTCTGGACAGAAAGA 2071
Qy 1948 GTCCAGGCCACGCTGCTGGTCTGCTGGCTTGGCCATGCTGCTGCTTGGCACA 2007
Db 2072 ATTCAAGTGTTCCTGGTAGTGTGCACTACTGTGTGTACCTTGGATGCTGCTGTTTAAA 2131
Qy 2008 CCCCTGCACCTGCTGCAACCGCCCGCCGCTCGGGAGGAGCCGCTGACCCGACAG 2067
Db 2132 CCATTTGGTCT-----TCGCCGTCAATTTTGGAGAGAAAGCAATTTGGGAATCTTC 2182
Qy 2068 GAGGAAACAGGCCGGTGTGGACCTGCTGACGCTGACGATCTGTAAATGGCTGGAGCTCC 2127
Db 2183 AACTTTGGTGGATCAGGGTGGGCAACCGACCGACAGAGGAGATCTGAGATTATTCAG 2242
Qy 2128 GATGAGGAAAGCGAGGGGCTGTGATGATGAAGAGAGCGGAGCTGCTGCCCTCCGAG 2187
Db 2243 CATGACCACTCTCCACCACCTCAGAGGACGACGAGTGTGACTTTGGGAGACCATG 2302
Qy 2188 GTGCTCATGCAACAGGCCATCCACCAATCGAGTTCTGCTGGGCTGGTCTTCCAACACC 2247

Db 2303 G-----TCCACCAAGCCATCCACACCATCGAGTACTGCCTGGGCTGCATCTCCAACACT 2356
Qy 2248 GCCTCTTACCTGCGCTCTGTGGGCCCTGAGCCTTGCCCCACGCCCAAGCTGTCCAGGTTCTG 2307
Db 2357 GCCTCTTACTTGGCGCTCTGGGCCCTCAGCCTTCGCTCATGCGCAGCTGTCTGAGGTGCTT 2416
Qy 2308 TGGGCCATGGTGATCGCATAGGCTTGGGCTTGGGCCCGGAGGTGGGCGTGGCGGCTG 2367
Db 2417 TGGACCATGGTGATCCACATCGGCTTGAGCGTGAAGAGCTTGGCCGG-----AGGTTTG 2470
Qy 2368 GTGCTGGTCCCCCATCTTTGCGCGCTTTGCGCTGTGATGACCGTGGCTATCTGCTGGTGATG 2427
Db 2471 GTGCTGTTCTTCTTCTTCACTGCTTGGCACCCCTGACCGTGGCCATCCTCCTGATCATG 2530
Qy 2428 GAGGACCTCTCAGCCTTCTGTCACGCCCTTGCGGCTGCACTGGGTGGAATTCAGAACAAAG 2487
Db 2531 GAGGGCCTCTCGGCCCTTTCTCCACGCACTGCGCTTACACTGGGTTGAGTTCCAGAATAAA 2590
Qy 2488 TTCTACTCAGGCACGGGCTACAAGCTGAGTCCCTTTCACCTTCG 2530
Db 2591 TTCTAAGGGGACCGGTTTCAAGTTCTTACCCTTCTCCTTCG 2633

Search completed: June 30, 2006, 07:42:25
Job time : 514 secs

| Result No. | Score | Query | | Length | DB | ID | Description |
|------------|--------|-------|------|--------|----------|----------|-------------|
| | | Match | % | | | | |
| 1 | 2640 | 100.0 | 2640 | 2 | AR212865 | AR212865 | Sequence |
| 2 | 2640 | 100.0 | 2640 | 2 | AR577091 | AR577091 | Sequence |
| 3 | 2640 | 100.0 | 2655 | 2 | AR270531 | AR270531 | Sequence |
| 4 | 2640 | 100.0 | 2655 | 2 | AR380371 | AR380371 | Sequence |
| 5 | 2640 | 100.0 | 2855 | 2 | AX332228 | AX332228 | Sequence |
| 6 | 2640 | 100.0 | 2855 | 2 | AX336173 | AX336173 | Sequence |
| 7 | 2640 | 100.0 | 2655 | 2 | AX410717 | AX410717 | Sequence |
| 8 | 2640 | 100.0 | 2655 | 5 | HSU45285 | HSU45285 | Human speci |
| 9 | 2597 | 98.4 | 2722 | 5 | BC032465 | BC032465 | Homo sapi |
| 10 | 2594 | 98.3 | 2691 | 5 | BC018133 | BC018133 | Homo sapi |
| 11 | 2550.2 | 96.6 | 2582 | 2 | CQ718250 | CQ718250 | Sequence |
| 12 | 2068 | 78.3 | 2488 | 2 | A98501 | A98501 | Sequence 1 |
| 13 | 2058 | 78.3 | 2488 | 2 | AX320850 | AX320850 | Sequence |
| 14 | 2068 | 78.3 | 2488 | 2 | BD080989 | BD080989 | Novel T c |
| 15 | 2068 | 78.3 | 2488 | 5 | AF025374 | AF025374 | Homo sapi |
| 16 | 2066.4 | 78.3 | 2488 | 2 | A98512 | A98512 | Sequence 12 |
| 17 | 2066.4 | 78.3 | 2488 | 2 | AX320861 | AX320861 | Sequence |
| 18 | 2066.4 | 78.3 | 2488 | 2 | BD080992 | BD080992 | Novel T c |

QY 241 GAGAGACCTTACCTTCTGAGGAGAGGTGCGCGGCTGGGTCTGCTGCCCG 300
DB 241 GAGAGACCTTACCTTCTGAGGAGAGGTGCGCGGCTGGGTCTGCTGCCCG 300
QY 301 CAAAGGGAGGTGCGGCAACCCCAACCCCGGACCTGCTGCGCATCCAGGAGGAGCG 360
DB 301 CAAAGGGAGGTGCGGCAACCCCAACCCCGGACCTGCTGCGCATCCAGGAGGAGCG 360
QY 361 GAGCGCTGCGCAGAGCTGCGGATGTCGGGCAACAGCAGGCGCTGCGGCGCAG 420
DB 361 GAGCGCTGCGCAGAGCTGCGGATGTCGGGCAACAGCAGGCGCTGCGGCGCAG 420
QY 421 CTGACACAGCTGAGCTCCAGCGCGCTGCTACGCCAGGCGCATGAACTCAGCTGGCA 480
DB 421 CTGACACAGCTGAGCTCCAGCGCGCTGCTACGCCAGGCGCATGAACTCAGCTGGCA 480
QY 481 GCGGCCCAACAGATGGGCGCTCAGAGGAGCGCCCTGCTCCAGGCGCCCGGGGCGG 540
DB 481 GCGGCCCAACAGATGGGCGCTCAGAGGAGCGCCCTGCTCCAGGCGCCCGGGGCGG 540
QY 541 CACAGGACCTGAGGCTCAACTTTGTCGAGGTGCGGTGGAGCCCAAGGCCCTGCG 600
DB 541 CACAGGACCTGAGGCTCAACTTTGTCGAGGTGCGGTGGAGCCCAAGGCCCTGCG 600
QY 601 CTAGAGCGCTGCTCTGAGGCGCTGCGCGGCTTCTCATTTGCCAGCTTCAAGGAGCTG 660
DB 601 CTAGAGCGCTGCTCTGAGGCGCTGCGCGGCTTCTCATTTGCCAGCTTCAAGGAGCTG 660
QY 661 GAGCAGCGCTGAGCAACCCCGTGCAGGCGAGCCAGCCACCTGAGTACCTTCTCATTC 720
DB 661 GAGCAGCGCTGAGCAACCCCGTGCAGGCGAGCCAGCCACCTGAGTACCTTCTCATTC 720
QY 721 TCTACTGGGTGAGCAGATCGGACAGAGATCCGCAAGATCAGGAGCTGCTTCCACTGC 780
DB 721 TCTACTGGGTGAGCAGATCGGACAGAGATCCGCAAGATCAGGAGCTGCTTCCACTGC 780
QY 781 CACGTCTTCCCTTCTGAGGAGGAGGCGCGCTCGGGCGCTGAGCAGCTGCAAA 840
DB 781 CACGTCTTCCCTTCTGAGGAGGAGGCGCGCTCGGGCGCTGAGCAGCTGCAAA 840
QY 841 GAGCAGCGAGGCTGAGAGGTCTCGGGGAGACAGAGGCTTCTGAGCAGGCTG 900
DB 841 GAGCAGCGAGGCTGAGAGGTCTCGGGGAGACAGAGGCTTCTGAGCAGGCTG 900
QY 901 CTAGGCGGCTGCTGAGCTGCTGCGCAGGCGAGGTGAGGTCCAAAGATGAAGGCC 960
DB 901 CTAGGCGGCTGCTGAGCTGCTGCGCAGGCGAGGTGAGGTCCAAAGATGAAGGCC 960
QY 961 GTGTACTGGCCCTGAACCAAGTGCAGCTGAGCAACAGCAAGTGCCTCATTTGCCGAG 1020
DB 961 GTGTACTGGCCCTGAACCAAGTGCAGCTGAGCAACAGCAAGTGCCTCATTTGCCGAG 1020
QY 1021 GCCTGGTGTCTGTGCGAGACTGCGCGCTGCGAGGAGGCGCTGCGGAGAGCTCGATG 1080
DB 1021 GCCTGGTGTCTGTGCGAGACTGCGCGCTGCGAGGAGGCGCTGCGGAGAGCTCGATG 1080
QY 1081 GAGGAGGAGTGAAGTGCCTGCTCAGCAGCTCCCTGCGGAGATGCCCCCAGACTC 1140
DB 1081 GAGGAGGAGTGAAGTGCCTGCTCAGCAGCTCCCTGCGGAGATGCCCCCAGACTC 1140
QY 1141 ATCCGCAACACCGTTCAAGGCGAGCTTCCAGGCGATGTTGAGTACGCGGTGGGC 1200
DB 1141 ATCCGCAACACCGTTCAAGGCGAGCTTCCAGGCGATGTTGAGTACGCGGTGGGC 1200
QY 1201 CGCTACGAGGAGTCAACCCCGCTGCTTACACCATCATCCTTCCCTTCTGTTGCT 1260
DB 1201 CGCTACGAGGAGTCAACCCCGCTGCTTACACCATCATCCTTCCCTTCTGTTGCT 1260
QY 1261 GTGATGTTGCGGAGTGGGCGACCGGCTGCTCATGTTCTTCTGCGCTGCGCAGTGC 1320
DB 1261 GTGATGTTGCGGAGTGGGCGACCGGCTGCTCATGTTCTTCTGCGCTGCGCAGTGC 1320
QY 1321 CTTGCGGAGAACCGACCGCTGTGAAAGCGCGCAGACGAGATCTGGCAGACTTTCTTC 1380
DB 1321 CTTGCGGAGAACCGACCGCTGTGAAAGCGCGCAGACGAGATCTGGCAGACTTTCTTC 1380
QY 1381 AGGGCCCGCTACTGCTCTGCTTTATGGGCTGTTCTCCATCTACACCGGCTTCACTAC 1440
DB 1381 AGGGCCCGCTACTGCTCTGCTTTATGGGCTGTTCTCCATCTACACCGGCTTCACTAC 1440
QY 1441 AACGAGTGTTCAGTTCGCGCCACAGCATCTTCCCTCGGGCTGGAGTGTGGCCGATG 1500
DB 1441 AACGAGTGTTCAGTTCGCGCCACAGCATCTTCCCTCGGGCTGGAGTGTGGCCGATG 1500
QY 1501 GCGCAACAGTCTGGCTGGAGTGCATCTTCTGGGCGAGCACACGATGCTTACCTGGAT 1560
DB 1501 GCGCAACAGTCTGGCTGGAGTGCATCTTCTGGGCGAGCACACGATGCTTACCTGGAT 1560
QY 1561 CCGCAACAGTCTGGCTGGAGTGCATCTTCTGGGCGAGCACACGATGCTTACCTGGAT 1620
DB 1561 CCGCAACAGTCTGGCTGGAGTGCATCTTCTGGGCGAGCACACGATGCTTACCTGGAT 1620
QY 1621 CTGGCTGCGCAACCTTGGCTTCTTCAAGTGAAGATGTCCTGATCTCTG 1680
DB 1621 CTGGCTGCGCAACCTTGGCTTCTTCAAGTGAAGATGTCCTGATCTCTG 1680
QY 1681 GCGCTGCTGCGCAACCTTGGCTTCTTCAAGTGAAGATGTCCTGATCTCTG 1740
DB 1681 GCGCTGCTGCGCAACCTTGGCTTCTTCAAGTGAAGATGTCCTGATCTCTG 1740
QY 1741 CAGAGGCGAGGCTGCTGCGGAGCGCTGCGGAGCTTCACTTCTGCTGGAGCTTTC 1800
DB 1741 CAGAGGCGAGGCTGCTGCGGAGCGCTGCGGAGCTTCACTTCTGCTGGAGCTTTC 1800
QY 1801 GGTACCTGCTGCTTCAAGTGAAGATGTCCTGATCTCTGCTGGAGCTTTC 1860
DB 1801 GGTACCTGCTGCTTCAAGTGAAGATGTCCTGATCTCTGCTGGAGCTTTC 1860
QY 1861 TCGCCAGCAGCTTCTTCAAGTGAAGATGTCCTGATCTCTGCTGGAGCTTTC 1920
DB 1861 TCGCCAGCAGCTTCTTCAAGTGAAGATGTCCTGATCTCTGCTGGAGCTTTC 1920
QY 1921 AGGCTGCTTACCCCGGAGGAGTGTGCGGAGCTTCAAGTGAAGATGTCCTGCTGGAGCTTTC 1980
DB 1921 AGGCTGCTTACCCCGGAGGAGTGTGCGGAGCTTCAAGTGAAGATGTCCTGCTGGAGCTTTC 1980
QY 1981 ATGCTGCTTACCCCGGAGGAGTGTGCGGAGCTTCAAGTGAAGATGTCCTGCTGGAGCTTTC 2040
DB 1981 ATGCTGCTTACCCCGGAGGAGTGTGCGGAGCTTCAAGTGAAGATGTCCTGCTGGAGCTTTC 2040
QY 2041 CTGCGGAGGAGGCTTCAAGTGAAGATGTCCTGATCTCTGCTGGAGCTTTC 2100
DB 2041 CTGCGGAGGAGGCTTCAAGTGAAGATGTCCTGATCTCTGCTGGAGCTTTC 2100
QY 2101 GACGATCTGTGAATGGCTGGAGTGTGCGGAGCTTCAAGTGAAGATGTCCTGCTGGAGCTTTC 2160
DB 2101 GACGATCTGTGAATGGCTGGAGTGTGCGGAGCTTCAAGTGAAGATGTCCTGCTGGAGCTTTC 2160
QY 2161 GAGGAGGAGGCTTCAAGTGAAGATGTCCTGATCTCTGCTGGAGCTTTC 2220
DB 2161 GAGGAGGAGGCTTCAAGTGAAGATGTCCTGATCTCTGCTGGAGCTTTC 2220
QY 2221 TTTCTGCTGGGCTGCTTCAAGTGAAGATGTCCTGATCTCTGCTGGAGCTTTC 2280
DB 2221 TTTCTGCTGGGCTGCTTCAAGTGAAGATGTCCTGATCTCTGCTGGAGCTTTC 2280
QY 2281 GCGCAGGAGGCTTCAAGTGAAGATGTCCTGATCTCTGCTGGAGCTTTC 2340
DB 2281 GCGCAGGAGGCTTCAAGTGAAGATGTCCTGATCTCTGCTGGAGCTTTC 2340
QY 2341 GCGCGGAGGCTTCAAGTGAAGATGTCCTGATCTCTGCTGGAGCTTTC 2400
DB 2341 GCGCGGAGGCTTCAAGTGAAGATGTCCTGATCTCTGCTGGAGCTTTC 2400
QY 2401 ATGACCGTGGCTATCTGCTGCTGATGAGGAGCTTCTCAGGCTTCTGCAAGCCTTGGG 2460
DB 2401 ATGACCGTGGCTATCTGCTGCTGATGAGGAGCTTCTCAGGCTTCTGCAAGCCTTGGG 2460

| | | | | |
|----|------|--|--|------|
| Db | 1561 | | CCCAACGTCACCGGTGCTCTCTCGGACCCCTACCCCTTTGGCATCGATTCGATATTTGGAGC | 1620 |
| Qy | 1621 | CTGGCTGCGCAACCACTTGAGCTTCTCAACTCCCTTCAAGATGAAGATGTCGGTCACTCCTG | 1680 | |
| Db | 1621 | CTGGCTGCCAACCACTTGAGCTTCTCAACTCCCTTCAAGATGAAGATGTCGGTCACTCCTG | 1680 | |
| Qy | 1681 | GGGTCGTFGCACATGGGCTTTGGGGTGGTCTCTCGAGTCTTCAAACCACTGCACTTTGGC | 1740 | |
| Db | 1681 | GGCGTCGTCACATAGGCTTTGGGGTGGTCTCTCGAGTCTTCAAACCACTGCACTTTGGC | 1740 | |
| Qy | 1741 | CAGAGGCACCGGCTGCTGGAGACGCTGCCGAGCTCACTTCCTGCTGGGACTCTTTC | 1800 | |
| Db | 1741 | CAGAGGCACCGGCTGCTGGAGACGCTGCCGAGCTCACTTCCTGCTGGGACTCTTTC | 1800 | |
| Qy | 1801 | GGTTACCTCTGTTTCTTAGTCACTACAAGTGGCTGTGTCTCTGGGCTGCCAGGGCGGCC | 1860 | |
| Db | 1801 | GGTTACCTCTGTTTCTTAGTCACTACAAGTGGCTGTGTCTCTGGGCTGCCAGGGCGGCC | 1860 | |
| Qy | 1861 | TCGCCCAGCATCTCATCCACTTCATCAACAATGTTCTTCTTCTCCACAGCCCCAGCAAC | 1920 | |
| Db | 1861 | TCGCCCAGCATCTCATCCACTTCATCAACAATGTTCTTCTTCTCCACAGCCCCAGCAAC | 1920 | |
| Qy | 1921 | AGGCTGCTCTACCCCGGACGAGGTGGTCCAGGCCACGCTGGTGGTCTCTGGGCTTTGGCC | 1980 | |
| Db | 1921 | AGGCTGCTCTACCCCGGACGAGGTGGTCCAGGCCACGCTGGTGGTCTCTGGGCTTTGGCC | 1980 | |
| Qy | 1981 | ATGSGTGCCCATCTCTGCTTGGGCACACCCCTGCACTGCTGCAACGCCACCCGCGCGCGC | 2040 | |
| Db | 1981 | ATGSGTGCCCATCTCTGCTTGGGCACACCCCTGCACTGCTGCAACGCCACCCGCGCGCGC | 2040 | |
| Qy | 2041 | CTGCGGAGGAGGCCCGCTGACCCGACAGGAGAAAAGGAGGAGGCTTGGGACCTTGCCT | 2100 | |
| Db | 2041 | CTGCGGAGGAGGCCCGCTGACCCGACAGGAGAAAAGGAGGAGGCTTGGGACCTTGCCT | 2100 | |
| Qy | 2101 | GACGCATCTGTGAATGCTGGAGCTCCGATGAGGAAAAGGAGGGGCTTGGATGATGAA | 2160 | |
| Db | 2101 | GACGCATCTGTGAATGCTGGAGCTCCGATGAGGAAAAGGAGGGGCTTGGATGATGAA | 2160 | |
| Qy | 2161 | GAGGAGGCCGAGCTCGTCCCTCCGAGGTGCTCATGCAACAGGCCATCCACACATCGAG | 2220 | |
| Db | 2161 | GAGGAGGCCGAGCTCGTCCCTCCGAGGTGCTCATGCAACAGGCCATCCACACATCGAG | 2220 | |
| Qy | 2221 | TTCTGCTTGGGCTCGGCTCTCAAACACCGCTCTTACTGTCGCTGTGGGCCCTTGAGCCTG | 2280 | |
| Db | 2221 | TTCTGCTTGGGCTCGGCTCTCAAACACCGCTCTTACTGTCGCTGTGGGCCCTTGAGCCTG | 2280 | |
| Qy | 2281 | GCCACGCCACAGCTGTCGAGGTTCTGCGGCCATGCTGATGCGCATAGGCTTGGGCTG | 2340 | |
| Db | 2281 | GCCACGCCACAGCTGTCGAGGTTCTGCGGCCATGCTGATGCGCATAGGCTTGGGCTG | 2340 | |
| Qy | 2341 | GGCCGGAGGTGGGCGTGGGCGCTGTGGTGTGCTGCCCATCTTTGCGCGCTTTGCGCGT | 2400 | |
| Db | 2341 | GGCCGGAGGTGGGCGTGGGCGCTGTGGTGTGCTGCCCATCTTTGCGCGCTTTGCGCGT | 2400 | |
| Qy | 2401 | ATGACCGTGGCTATCTCTGCTGGTGAATGAGAGGACTCTCAGCGCTTCTGCAAGCCCTGCGG | 2460 | |
| Db | 2401 | ATGACCGTGGCTATCTCTGCTGGTGAATGAGAGGACTCTCAGCGCTTCTGCAAGCCCTGCGG | 2460 | |
| Qy | 2461 | CTGCACCTGGGTGGAAATTCAGAACAGTTCTACTCAGGCAAGGCTACAGCTGAGTCCC | 2520 | |
| Db | 2461 | CTGCACCTGGGTGGAAATTCAGAACAGTTCTACTCAGGCAAGGCTACAGCTGAGTCCC | 2520 | |
| Qy | 2521 | TTCACTTTCGTCGCCACAGATGACTAGGGGCCACTGCAAGTCTCTGCCAGACCTCTTCTCT | 2580 | |
| Db | 2521 | TTCACTTTCGTCGCCACAGATGACTAGGGGCCACTGCAAGTCTCTGCCAGACCTCTTCTCT | 2580 | |
| Qy | 2581 | GACCTCTGAGGCAGGAGGAAATGAAGACGGTCCGCTCGGCAAAAAAATAAAAAA | 2640 | |
| Db | 2581 | GACCTCTGAGGCAGGAGGAAATGAAGACGGTCCGCTCGGCAAAAAAATAAAAAA | 2640 | |

QY 241 GAGAGACCTTACCTTCTGCGAGGAGGTGGCGGGCTGGGTGGTCTGTCGCCCGC 300
DB 241 GAGAAGACCTTACCTTCTGCGAGGAGGAGGTGGCGGGCTGGGTGGTCTGTCGCCCGC 300
QY 301 CCAAGAGGAGGCTGCGCGACCCACCCCGGAGCCTGTGCGCATCCAGGAGGAGCG 360
DB 301 CCAAGAGGAGGCTGCGCGACCCACCCCGGAGCCTGTGCGCATCCAGGAGGAGCG 360
QY 361 GAGCGCTGCGCCAGGAGCTGCGGATGTGCGGGCAACAGAGGCCCTGCGGGCCAG 420
DB 361 GAGCGCTGCGCCAGGAGCTGCGGATGTGCGGGCAACAGAGGCCCTGCGGGCCAG 420
QY 421 CTGCAACAGCTGAGCTTCCAGCGCGCGCTGCTAGCGCCAGGCGCATGAACTCAGCTGGCA 480
DB 421 CTGCAACAGCTGAGCTTCCAGCGCGCGCTGCTAGCGCCAGGCGCATGAACTCAGCTGGCA 480
QY 481 GCGCGCCACACAGATGGGGCTCAGAGAGAGCGCCCTGTCCAGGCGCCCGGGGGCGG 540
DB 481 GCGCGCCACACAGATGGGGCTCAGAGAGAGCGCCCTGTCCAGGCGCCCGGGGGCGG 540
QY 541 CACAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGGTGAGCGCCCAAGGCCCTTGCC 600
DB 541 CACAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGGTGAGCGCCCAAGGCCCTTGCC 600
QY 601 CTAGAGGCGCTGCTGAGGGCTGCGCGGGCTTCTCATTTGCGCAGCTTCAGGAGCTG 660
DB 601 CTAGAGGCGCTGCTGAGGGCTGCGCGGGCTTCTCATTTGCGCAGCTTCAGGAGCTG 660
QY 661 GAGCAGCGCTGAGCAGCCCGTACGCGGCGAGCAGCAGCAGCTGGATGACTTCTTCATC 720
DB 661 GAGCAGCGCTGAGCAGCCCGTACGCGGCGAGCAGCAGCAGCTGGATGACTTCTTCATC 720
QY 721 TCTACTGGGGTGAGCAGATCGGACAGAAATCGCAAGATCACGGACTGCTTCCACTGC 780
DB 721 TCTACTGGGGTGAGCAGATCGGACAGAAATCGCAAGATCACGGACTGCTTCCACTGC 780
QY 781 CAGCTTCTCCGTTTCTGAGCAGGAGGAGCGCGCTCGGGCCCTGAGCAGCTGCA 840
DB 781 CAGCTTCTCCGTTTCTGAGCAGGAGGAGCGCGCTCGGGCCCTGAGCAGCTGCA 840
QY 841 CAGCAGGCGAGGCTGAGGAGGTCTCGGGGAGACAGCGGTTCCTGAGCAGGTG 900
DB 841 CAGCAGGCGAGGCTGAGGAGGTCTCGGGGAGACAGCGGTTCCTGAGCAGGTG 900
QY 901 CTAGCGGGTGCTGAGCTGCTGCGCGCAGGCGAGGTGCAAGATGAAGGCC 960
DB 901 CTAGCGGGTGCTGAGCTGCTGCGCGCAGGCGAGGTGCAAGATGAAGGCC 960
QY 961 GTGTACTGGCCCTGAACAGTGCAGGTGAGCAGCAGCAGTGCCTCATTTGCCGAG 1020
DB 961 GTGTACTGGCCCTGAACAGTGCAGGTGAGCAGCAGCAGTGCCTCATTTGCCGAG 1020
QY 1021 GCCTGTGTCTGTGAGACCTGCGCGCTCAGAGGCGCTGCGGGACAGCTCGATG 1080
DB 1021 GCCTGTGTCTGTGAGACCTGCGCGCTCAGAGGCGCTGCGGGACAGCTCGATG 1080
QY 1081 GAGGAGGAGTGAAGTGGGTGCTACCGCATCCCTGCGGGACATGCCCCCACATC 1140
DB 1081 GAGGAGGAGTGAAGTGGGTGCTACCGCATCCCTGCGGGACATGCCCCCACATC 1140
QY 1141 ATCCGACACACCGCTTACGGCCAGTTCCAGGGCATCTGGATCGCTACGGCTGGGC 1200
DB 1141 ATCCGACACACCGCTTACGGCCAGTTCCAGGGCATCTGGATCGCTACGGCTGGGC 1200
QY 1201 CGCTACAGAGGTCAACCCCGCTCCTACCATCATCATCTCCCTTCCCTTGTGTCT 1260
DB 1201 CGCTACAGAGGTCAACCCCGCTCCTACCATCATCATCTCCCTTCCCTTGTGTCT 1260
QY 1261 GTGATGTTGGGATGTGGGCGACGCGCTGCTCATGTTCTTTCGCGCTTGGCCATGTC 1320
DB 1261 GTGATGTTGGGATGTGGGCGACGCGCTGCTCATGTTCTTTCGCGCTTGGCCATGTC 1320

QY 1321 CTTGCGGAGAACCGACCGGCTGTAAAGCCGCGAGAAAGAGATCTGGCAGACTTCTTC 1380
DB 1321 CTTGCGGAGAACCGACCGGCTGTAAAGCCGCGAGAAAGAGATCTGGCAGACTTCTTC 1380
QY 1381 AGGGGCGCTTACCTGCTCCTGCTTATGGGCTGTTCCTCATCTACACCGGCTTCACTAC 1440
DB 1381 AGGGGCGCTTACCTGCTCCTGCTTATGGGCTGTTCCTCATCTACACCGGCTTCACTAC 1440
QY 1441 AACGAGTGTCTTCACTGCGCCACAGCATCTTCCCTCGGGCTGGAGTGTGGCCCATG 1500
DB 1441 AACGAGTGTCTTCACTGCGCCACAGCATCTTCCCTCGGGCTGGAGTGTGGCCCATG 1500
QY 1501 GCCAACGAGTGTGGCTGAGTGAATTCCTGGCCAGCAGCAGATGCTTACCTTGAT 1560
DB 1501 GCCAACGAGTGTGGCTGAGTGAATTCCTGGCCAGCAGCAGATGCTTACCTTGAT 1560
QY 1561 CCCAACGCTACCGGTGTCTTCTGGGACCTACCCCTTTGGCATCGATCTTATTTGAGC 1620
DB 1561 CCCAACGCTACCGGTGTCTTCTGGGACCTACCCCTTTGGCATCGATCTTATTTGAGC 1620
QY 1621 CTGGCTGCCAACCACTTGAAGTCTTCTCAACTCTTCAAGATGAAGATGCTCCGTCACTG 1680
DB 1621 CTGGCTGCCAACCACTTGAAGTCTTCTCAACTCTTCAAGATGAAGATGCTCCGTCACTG 1680
QY 1681 GCGCTGTCACATGCGCTTTGGGGTGGTCTCGGAGTCTTCAACAGCTGCATTTGGC 1740
DB 1681 GCGCTGTCACATGCGCTTTGGGGTGGTCTCGGAGTCTTCAACAGCTGCATTTGGC 1740
QY 1741 CAGAGGCAACCGGTGTCTGCGGAGCGCTCCCGAGCTCACCTTCTGCTGGGACTTCTC 1800
DB 1741 CAGAGGCAACCGGTGTCTGCGGAGCGCTCCCGAGCTCACCTTCTGCTGGGACTTCTC 1800
QY 1801 GGTACTCTGCTTCTAGTCACTCAAGTGGCTGTGTCTGCTGGCTGCGAGCGCGCC 1860
DB 1801 GGTACTCTGCTTCTAGTCACTCAAGTGGCTGTGTCTGCTGGCTGCGAGCGCGCC 1860
QY 1861 TCGCCAGCATCTCATCCACTTTCATCAAGTGGTCTTCTTCCACAGCGCCAGCAAC 1920
DB 1861 TCGCCAGCATCTCATCCACTTTCATCAAGTGGTCTTCTTCCACAGCGCCAGCAAC 1920
QY 1921 AGGCTGCTTACCCCGGAGGAGGTGTCAGGCGACAGCTGTGTGCTTCTGGCTTGGCC 1980
DB 1921 AGGCTGCTTACCCCGGAGGAGGTGTCAGGCGACAGCTGTGTGCTTCTGGCTTGGCC 1980
QY 1981 ATGGTCCCATCTGCTGTGTCACACCCCTGCACTGTGCAACCGCCACCGCGCGCC 2040
DB 1981 ATGGTCCCATCTGCTGTGTCACACCCCTGCACTGTGCAACCGCCACCGCGCGCC 2040
QY 2041 CTGCGGAGGAGCGCGCTGACCGACAGGAGGAAAACAGGCGGGTGTGCTGGACCTGCT 2100
DB 2041 CTGCGGAGGAGCGCGCTGACCGACAGGAGGAAAACAGGCGGGTGTGCTGGACCTGCT 2100
QY 2101 GAGCATCTGTGAATGGCTGAGCTCCGATGAGAAAAGGAGGGGCTTGGATGATGA 2160
DB 2101 GAGCATCTGTGAATGGCTGAGCTCCGATGAGAAAAGGAGGGGCTTGGATGATGA 2160
QY 2161 GAGGAGGCGAGCTGCTCCCTCCGAGGTGCTCATGCAACAGGCGCATCCACCATCGAG 2220
DB 2161 GAGGAGGCGAGCTGCTCCCTCCGAGGTGCTCATGCAACAGGCGCATCCACCATCGAG 2220
QY 2221 TTTGCTGCTGGCTGCTGCTTCCAAACACCGCTCTCACTGCGCTGTGGGCGCTGAGCCTG 2280
DB 2221 TTTGCTGCTGGCTGCTGCTTCCAAACACCGCTCTCACTGCGCTGTGGGCGCTGAGCCTG 2280
QY 2281 GCCACGCGCAGCTGCTGGGCTTCTGTGGGCTGCTGTGATGCGCATAGGCTTGGGCTG 2340
DB 2281 GCCACGCGCAGCTGCTGGGCTTCTGTGGGCTGCTGTGATGCGCATAGGCTTGGGCTG 2340
QY 2341 GGCAGGAGGTGGGCTGCGGCTGTGGTCCCATCTTTGCGCTTTGCGCTG 2400
DB 2341 GGCAGGAGGTGGGCTGCGGCTGTGGTCCCATCTTTGCGCTTTGCGCTG 2400
QY 2401 ATGACCGTGGCTATCTGCTGTGTGATGAGGAGTCTCAGCCTTCTGCGAGCGCTGCGG 2460

```
Db 2401 ATGACCGTGGCTATCTGCTGGTGTGGAGGACTCTCAGCCTTCTGCAAGCCCTGCGG 2460
Qy 2461 CTGCACTGGGTGGAATTCAGAACAAAGTTCTACTCAGGCAAGGGCTTACAAAGCTGAGTCCC 2520
Db 2461 CTGCACTGGGTGGAATTCAGAACAAAGTTCTACTCAGGCAAGGGCTTACAAAGCTGAGTCCC 2520
Qy 2521 TTCACTCTCGCTGCCACAGATGACTAGGGCCCACTGCAGGTCTGCGCAGACCTCTTCTTCT 2580
Db 2521 TTCACTCTCGCTGCCACAGATGACTAGGGCCCACTGCAGGTCTGCGCAGACCTCTTCTTCT 2580
Qy 2581 GACCTCTGAGGAGGAGAGGAATAAGACGGTCCGCGCTTGGCAAAAAAAGAAAAA 2640
Db 2581 GACCTCTGAGGAGGAGAGGAATAAGACGGTCCGCGCTTGGCAAAAAAAGAAAAA 2640

RESULT 7
AX410717
LOCUS AX410717 2655 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 3364 from Patent WO0229103.
ACCESSION AX410717
VERSION AX410717.1 GI:21443422
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE
1.
AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 3364 11-APR-2002;
GENE LOGIC INC. (US)
FEATURES
source
location/Qualifiers
1..2655
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. U45285"

ORIGIN
Query Match 100.0%; Score 2640; DB 2; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGCGTGGCGGACCGGACGAGCAGCGAGCGGCGGCGGAGCACACCCGGGGACCATG 60
Db 1 CGGCGTGGCGGACCGGACGAGCAGCGAGCGGCGGCGGAGCACACCCGGGGACCATG 60
Qy 61 GGCTCCATGTTCCGGAGGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCACAGGGCT 120
Db 61 GGCTCCATGTTCCGGAGGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCACAGGGCT 120
Qy 121 GCCTACACCTGCTGAGTGGCTGGGCGAGCTGGGCGCTCGTGGAGTTTACAGACCTCAAC 180
Db 121 GCCTACACCTGCTGAGTGGCTGGGCGAGCTGGGCGCTCGTGGAGTTTACAGACCTCAAC 180
Qy 181 GCCTGGTGGAGCGCTTCAGAGACGCTTTGTGGTTGATGTTGGCGCTGTAGAGGCTG 240
Db 181 GCCTGGTGGAGCGCTTCAGAGACGCTTTGTGGTTGATGTTGGCGCTGTAGAGGCTG 240
Qy 241 GAGAGACCTTCACCTTCTGAGGAGGAGGTGGCGGCGTGGCTGGCTGGCTGGCCCG 300
Db 241 GAGAGACCTTCACCTTCTGAGGAGGAGGTGGCGGCGTGGCTGGCTGGCTGGCCCG 300
Qy 301 CAAAAGGGAGGCTCGCGGACACCCCAACCCCGGACCTGCTGCGATCCAGGAGGAGAG 360
Db 301 CAAAAGGGAGGCTCGCGGACACCCCAACCCCGGACCTGCTGCGATCCAGGAGGAGAG 360
Qy 361 GAGCGCTGGCCAGAGCTGGGATGTGGGGGCAACACAGCGCCCTGCGGGCCG 420
Db 361 GAGCGCTGGCCAGAGCTGGGATGTGGGGGCAACACAGCGCCCTGCGGGCCG 420
```

```
Qy 421 CTGACACAGCTGACAGCTCCACGCGCGCTGCTACGCGAGGGCCATGAACCTCAGCTGGCA 480
Db 421 CTGACACAGCTGACAGCTCCACGCGCGCTGCTACGCGAGGGCCATGAACCTCAGCTGGCA 480
Qy 481 GCGCGCCACACAGATGGGGCTCAGAGGAGCGCCCTGCTCAGGCGCCCGGGGGCGG 540
Db 481 GCGCGCCACACAGATGGGGCTCAGAGGAGCGCCCTGCTCAGGCGCCCGGGGGCGG 540
Qy 541 CACACGACCTGAGGGTCAACTTTGTGCAAGTGCCTGAGGAGCCCAAGGCGCCCTGCC 600
Db 541 CACACGACCTGAGGGTCAACTTTGTGCAAGTGCCTGAGGAGCCCAAGGCGCCCTGCC 600
Qy 601 CTAGAGCGCTGCTCTGAGGGGCTTCCGCGCGCTTCTCATTTGCCAGCTTCAGGAGCTG 660
Db 601 CTAGAGCGCTGCTCTGAGGGGCTTCCGCGCGCTTCTCATTTGCCAGCTTCAGGAGCTG 660
Qy 661 GAGCAGCGCTGGAGCACCCCGTGACGGCGAGCCAGCCACGTGATGATCTTCTTCATC 720
Db 661 GAGCAGCGCTGGAGCACCCCGTGACGGCGAGCCAGCCACGTGATGATCTTCTTCATC 720
Qy 721 TCCTACTGGGTTGACAGATCGGACAGAGATCCGCAAGATCACCGACTGCTTCCACTGC 780
Db 721 TCCTACTGGGTTGACAGATCGGACAGAGATCCGCAAGATCACCGACTGCTTCCACTGC 780
Qy 781 CACGTCTTCCCGTTTCTGACGAGGAGGCGCCCTCGGGGCGCTGCGAGCTGCAAA 840
Db 781 CACGTCTTCCCGTTTCTGACGAGGAGGCGCCCTCGGGGCGCTGCGAGCTGCAAA 840
Qy 841 CAGCAGACGAGGAGCTGACGAGGTCTCGGGGAGACAGAGCGTCTCTGAGCCAGGTG 900
Db 841 CAGCAGACGAGGAGCTGACGAGGTCTCGGGGAGACAGAGCGTCTCTGAGCCAGGTG 900
Qy 901 CTAGGCGGGTGTCTGAGTGTGCGGACAGGAGGTGCGAGGTCCACAAGATGAAGGC 960
Db 901 CTAGGCGGGTGTCTGAGTGTGCGGACAGGAGGTGCGAGGTCCACAAGATGAAGGC 960
Qy 961 GTGTACTGGCCCTGAACAGTGCAGGTGAGCACACGCAAGTGCCTCATTCGCGAG 1020
Db 961 GTGTACTGGCCCTGAACAGTGCAGGTGAGCACACGCAAGTGCCTCATTCGCGAG 1020
Qy 1021 GCCTGGTGTCTGTGCGAGACCTGCGCGCTGCGAGGAGCGCTCGGGACAGCTCGATG 1080
Db 1021 GCCTGGTGTCTGTGCGAGACCTGCGCGCTGCGAGGAGCGCTCGGGACAGCTCGATG 1080
Qy 1081 GAGGAGGAGTGTGCGTGTGCTCAGCGATCCCTGCGCGGGAATGCCCCCACAATC 1140
Db 1081 GAGGAGGAGTGTGCGTGTGCTCAGCGATCCCTGCGCGGGAATGCCCCCACAATC 1140
Qy 1141 ATCCGCAACCAACCGCTTCAAGGCGAGCTTCCAGGGCATCGTGGATCGCTACGGCGTGGC 1200
Db 1141 ATCCGCAACCAACCGCTTCAAGGCGAGCTTCCAGGGCATCGTGGATCGCTACGGCGTGGC 1200
Qy 1201 CGCTACAGGAGTCAACCCCGCTTCTACACCATCATCCTTCCCTTCTGTTTGTCT 1260
Db 1201 CGCTACAGGAGTCAACCCCGCTTCTACACCATCATCCTTCCCTTCTGTTTGTCT 1260
Qy 1261 GTGATGTTGGGGATGTGGGCAAGGGCTGTCTGCTTCTTCCCTGCGCGATGCTC 1320
Db 1261 GTGATGTTGGGGATGTGGGCAAGGGCTGTCTGCTTCTTCCCTGCGCGATGCTC 1320
Qy 1321 CTTGGGAGAAACCGACCGGCTGTAAAGCCGCGAGAAACAGATCTGGCAGACTTTCTTC 1380
Db 1321 CTTGGGAGAAACCGACCGGCTGTAAAGCCGCGAGAAACAGATCTGGCAGACTTTCTTC 1380
Qy 1381 AGGGCGCTACCTGCTCTCTTATGGGCTTGTTCATCTACACCGGCTTCTATCTAC 1440
Db 1381 AGGGCGCTACCTGCTCTCTTATGGGCTTGTTCATCTACACCGGCTTCTATCTAC 1440
Qy 1441 AACGAGTCTTCAGTGGCGCACAGCATCTTCCCTCGGGCTGAGGTGCGCGCCATG 1500
Db 1441 AACGAGTCTTCAGTGGCGCACAGCATCTTCCCTCGGGCTGAGGTGCGCGCCATG 1500
Qy 1501 GCCAACAGTCTGGGTGAGTGTGCAATTCCTGGGCCAGCACACAGATGCTTACCTGGAT 1560
```

1501 GCGAAGCTCTGGTGGAGTATGCAATCTGGCCAGCACGATGCTTACCTGGAT 1560
1561 CCCAAGCTACCGGTGCTTCTGGGACCTACCCCTTTGGCATCGATCTATTTGGAGC 1620
1561 CCCAAGCTACCGGTGCTTCTGGGACCTACCCCTTTGGCATCGATCTATTTGGAGC 1620
1621 CTGGCTGCCAACCACTTGGCTTCTCAACTCTTCAAGATGAAGATGTCGCTCATCTG 1680
1621 CTGGCTGCCAACCACTTGGCTTCTCAACTCTTCAAGATGAAGATGTCGCTCATCTG 1680
1681 GCGCTCGTGACATGGCTTTGGGTGGTCTCTCGAGTCTTCAACACGCTGCACTTTGGC 1740
1681 GCGCTCGTGACATGGCTTTGGGTGGTCTCTCGAGTCTTCAACACGCTGCACTTTGGC 1740
1741 CAGAGGACCGGTGCTGCTGGAGAGCTGCCGAGCTCACCTTCTGCTGGGACTTCTC 1800
1741 CAGAGGACCGGTGCTGCTGGAGAGCTGCCGAGCTCACCTTCTGCTGGGACTTCTC 1800
1801 GGTACTCTGTGCTTCTAGTCACTTCAAGTGGCTGTGTCTGGGTGCTGAGGGCGCC 1860
1801 GGTACTCTGTGCTTCTAGTCACTTCAAGTGGCTGTGTCTGGGTGCTGAGGGCGCC 1860
1861 TCGCCAGCATCTCATCTCACTTCAACATGTTCTCTTCTCCACAGCCCGACCAAC 1920
1861 TCGCCAGCATCTCATCTCACTTCAACATGTTCTCTTCTCCACAGCCCGACCAAC 1920
1921 AGGCTGCTTACCCCGGAGAGGTGCTCCAGGACCACTGCTGTGTCTTGGCTTGGCC 1980
1921 AGGCTGCTTACCCCGGAGAGGTGCTCCAGGACCACTGCTGTGTCTTGGCTTGGCC 1980
1981 ATGGTGGCATCTGCTGTGCTGACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
1981 ATGGTGGCATCTGCTGTGCTGACACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
2041 CTGGGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
2041 CTGGGAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
2101 GAGCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGGAGGAGGAGGAGGAGGAGG 2160
2101 GAGCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGGAGGAGGAGGAGGAGGAGG 2160
2161 GAGGAGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
2161 GAGGAGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
2221 TTCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
2221 TTCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
2281 GCCACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
2281 GCCACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
2341 GCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
2341 GCGCGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
2401 ATGACCGTGGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
2401 ATGACCGTGGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
2461 CTGCACTGGGTGGAATTCAGAACAAAGTCTTACTCAGGACCGGCTTACAGCTGAGTCCC 2520
2461 CTGCACTGGGTGGAATTCAGAACAAAGTCTTACTCAGGACCGGCTTACAGCTGAGTCCC 2520
2521 TTACCTTCTGCTGCAAGATGCTAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
2521 TTACCTTCTGCTGCAAGATGCTAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
2581 GACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640

2581 GACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640

RESULT 8
HSU45285
LOCUS
DEFINITION
mRNA, complete cds.
ACCESSION
U45285
KEYWORDS
U45285.1 GI:1245045
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE
1 (bases 1 to 2655)
Li, Y.-P., Chen, W. and Staehenko, P.
Molecular cloning and characterization of a putative novel human
osteoclast-specific 116-kDa vacuolar proton pump subunit
Biochem. Biophys. Res. Commun. 218 (3), 813-821 (1996).
REFERENCE
2 (bases 1 to 2655)
Li, Y.-P., Chen, W. and Staehenko, P.
Direct Submission
Submitted (09-JAN-1996) Yi-Ping Li, Cytokine Biology, Forsyth
Dental Center, 140 Fenway, Boston, MA 02115, USA
LOCATION/Qualifiers
1. .2655
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue type="osteoclastoma tumor"
1. .2655
/gene="OC-116kDa"
58. 2547
/gene="OC-116kDa"
/note="novel membrane bound protein with at least six
transmembrane domains"
/codon start=1
/product="specific 116-kDa vacuolar proton pump subunit"
/protein_id="AAA97878.1"
/db_xref="GI:1245046"
/translation="MGSMRSEEVALLVQLPTAAAYTCVSRIGELGLVEFRDLNASV
SAPQRVVDVVRCEELEKFTTFLQEVRRAGLVLPKGRIPAPPDRILLRIQEEETE
RFAQLRDVNRQALRAQLHQLHAAVLRQHEPOLAAHTDGAERTPLQAPCG
PHQDLRVNFAVAGVPHKAPALERLLWRACRGLIASPRELOPLSHPTGEPATWMT
FLISWGEIQGKIRKIDTCFCHVPPFLQBEARLQALQOQOQOELQVLETER
FUSQVLRVQLLPPGQVQVHKMAVYLALNOCVSSTTHKCLIAEAWCSVDLPALQOE
ALRDSMEGVSVAHRIPCRDMPPTLIETNRFTASFGQIVDRYGVGYQEVNPAFY
IITFFLFAVMFGDVGHLLMFLFALAVLAEANPAVKAQNEIWTQTFPRGYLLLM
LGPIYPTGPIYNECFSRATSIFFSGWSVAAMQSGWDAFLAQHTMLDLPNTGVF
LLETLPBELTFLGLFGLVFLVIYKWLVCWAARAASPSLIHFINFLSHSPNRLL
YPRQVQATLVVLAALMPLILGLTPHLHRRRLRRRRADROENKAGLLDLPD
ASVNSWDEEKAGLLDDEEAEVPEVLMQHTIETFCIGCVSNTASYRLSLALS
LAHAQLSEVLMVNRIGLGLREVGVAAVLPVIPAFAVMTVAILLVMEGLSFLH
ALRLHWVFQNKFSYGTGYKLSPTFAATDD"

ORIGIN
Query Match 100.0%; Score 2640; DB 5; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGCGTGGCGGACCGGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
DB 1 CCGCGTGGCGGACCGGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
QY 61 GGTCTCATGTTCCGGAGCGAGGAGTGGCTTCCAGCTCTTTCGCCCACAGCGGCT 120
DB 61 GGTCTCATGTTCCGGAGCGAGGAGTGGCTTCCAGCTCTTTCGCCCACAGCGGCT 120
QY 121 GCTACACCTCGTGGTGGCTGGCGGAGCTGGCGGCTCGTGGAGTTTCAGACCTCAAC 180

|||||
Db 121 GCCTACACCTGCGTGAGTCGGCTGGCGAGCTGGGCTCTGTGGAGTTTCAGAGACCTCAAC 180
QY 181 GCGTCGGTGAGCGCCTTCAGAGACGCTTTGTGGTTGATGTTTGGCGCTGTGAGAGCTG 240
Db 181 GCGTCGGTGAGCGCCTTCAGAGACGCTTTGTGGTTGATGTTTGGCGCTGTGAGAGCTG 240
QY 241 GAGAGACCTTCACCTTCCTGAGAGAGAGGTGCGCGGGCTGGGCTGGTTCCTGCCCCG 300
Db 241 GAGAGACCTTCACCTTCCTGAGAGAGAGGTGCGCGGGCTGGGCTGGTTCCTGCCCCG 300
QY 301 CAAAGGGAGGCTGCGGCAACCCCAACCCCGGACCTGCTGCGCATCCAGGAGGAGAG 360
Db 301 CAAAGGGAGGCTGCGGCAACCCCAACCCCGGACCTGCTGCGCATCCAGGAGGAGAG 360
QY 361 GAGCGCTGGCCAGAGCTGGGGATGTGCGGGGCAACAGAGGCGCTGCGGGCCAG 420
Db 361 GAGCGCTGGCCAGAGCTGGGGATGTGCGGGGCAACAGAGGCGCTGCGGGCCAG 420
QY 421 CTGCAACAGCTGAGCTCCACGCGCGTGCTACGCCAGGGCCATGAACTTCAGCTGGCA 480
Db 421 CTGCAACAGCTGAGCTCCACGCGCGTGCTACGCCAGGGCCATGAACTTCAGCTGGCA 480
QY 481 GCGGCCACACAGATGGGCTCAGAGAGGACGCCCCCTGCTCCAGGCCCCCGGGGGCG 540
Db 481 GCGGCCACACAGATGGGCTCAGAGAGGACGCCCCCTGCTCCAGGCCCCCGGGGGCG 540
QY 541 CACAGGACCTGAGGTTCAACTTTGTGGAGGTGCGCGGGCAACAGAGGCGCTGCGG 600
Db 541 CACAGGACCTGAGGTTCAACTTTGTGGAGGTGCGCGGGCAACAGAGGCGCTGCGG 600
QY 601 CTAGAGCGCTGCTCTGGAGGCGCTGCGCGGCTTCTCATTTGCCAGCTTCAGGAGCTG 660
Db 601 CTAGAGCGCTGCTCTGGAGGCGCTGCGCGGCTTCTCATTTGCCAGCTTCAGGAGCTG 660
QY 661 GAGCAGCGCTGAGCACCCTGTGACGGCGAGCCAGCCAGCTGGATGATCTTCCTCATC 720
Db 661 GAGCAGCGCTGAGCACCCTGTGACGGCGAGCCAGCCAGCTGGATGATCTTCCTCATC 720
QY 721 TCTACTGGGTTAGCAGATCGACAGAGATCCGAGATCCGAGATCACGGACTGCTTCACCTGC 780
Db 721 TCTACTGGGTTAGCAGATCGACAGAGATCCGAGATCCGAGATCACGGACTGCTTCACCTGC 780
QY 781 CACGTCCTCCGCTTCTGACAGAGAGGCGCGCTCGGGCGCTGACAGCTGCA 840
Db 781 CACGTCCTCCGCTTCTGACAGAGAGGCGCGCTCGGGCGCTGACAGCTGCA 840
QY 841 GAGCAGCGAGGCTGAGAGGTCTCTGGGGAGACAGAGCGGTTCCTGAGCAGGTG 900
Db 841 GAGCAGCGAGGCTGAGAGGTCTCTGGGGAGACAGAGCGGTTCCTGAGCAGGTG 900
QY 901 CTAGGCGGGTGTGAGCTGCTGCGCCAGGGCAGGTGCAAGTGAAGGCC 960
Db 901 CTAGGCGGGTGTGAGCTGCTGCGCCAGGGCAGGTGCAAGTGAAGGCC 960
QY 961 GTGTACTGGGCTGAAACAGTGCAGCTGAGCACAACGCAAGTGCCTCATTTGCCGAG 1020
Db 961 GTGTACTGGGCTGAAACAGTGCAGCTGAGCACAACGCAAGTGCCTCATTTGCCGAG 1020
QY 1021 GCGTGGTGTCTGTGGAGACTGCGCGCTGACAGGAGGCGCTGCGGACAGCTCGATG 1080
Db 1021 GCGTGGTGTCTGTGGAGACTGCGCGCTGACAGGAGGCGCTGCGGACAGCTCGATG 1080
QY 1081 GAGGAGGAGTGAAGTGGCTGACCGCATCCCTGCGGGAGATGCCCCCACTC 1140
Db 1081 GAGGAGGAGTGAAGTGGCTGACCGCATCCCTGCGGGAGATGCCCCCACTC 1140
QY 1141 ATCCGACCAACCGCTTCACGCGCAGCTTCAGAGGCACTCGTGGATCGCTACGGCTGGGC 1200
Db 1141 ATCCGACCAACCGCTTCACGCGCAGCTTCAGAGGCACTCGTGGATCGCTACGGCTGGGC 1200
QY 1201 CGCTACAGGAGGTCAACCCCGCTCCCTACACATCATCACTTCCTCCCTTCTGTTGCT 1260
|||||

Db 1201 CGCTACAGGAGGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCTCTGTTGCT 1260
QY 1261 GTGATGTTGCGGGATGTGGGCCACCGGCTGTCTATGTTCTCTTTCGCCCTTGGCCATGCTC 1320
Db 1261 GTGATGTTGCGGGATGTGGGCCACCGGCTGTCTATGTTCTCTTTCGCCCTTGGCCATGCTC 1320
QY 1321 CTTTGGGAGAACCGACCGGCTGTGAAAGCGCGGAGAACGAGATCTGGCAGACTTTCTTC 1380
Db 1321 CTTTGGGAGAACCGACCGGCTGTGAAAGCGCGGAGAACGAGATCTGGCAGACTTTCTTC 1380
QY 1381 AGGGGCGCTACTGCTCTCTGCTATATGGGCTGTTCTCCATCTACACCGGCTTCATCTAC 1440
Db 1381 AGGGGCGCTACTGCTCTCTGCTATATGGGCTGTTCTCCATCTACACCGGCTTCATCTAC 1440
QY 1441 AACGAGTGTTCAGTGGGCCACACAGCATCTTCCCTCGGGCTGAGTGTGGCCGCGATG 1500
Db 1441 AACGAGTGTTCAGTGGGCCACACAGCATCTTCCCTCGGGCTGAGTGTGGCCGCGATG 1500
QY 1501 GCCAACAGTCTGGCTGGAGTGATGCACTTCTGGGCCACAGCACAGATGCTTACCCCTGGAT 1560
Db 1501 GCCAACAGTCTGGCTGGAGTGATGCACTTCTGGGCCACAGCACAGATGCTTACCCCTGGAT 1560
QY 1561 CCCAACGTCAACGGTGTCTCTGGGACCCCTAACCTTTGGCATCGATCTTATTTGGAGC 1620
Db 1561 CCCAACGTCAACGGTGTCTCTGGGACCCCTAACCTTTGGCATCGATCTTATTTGGAGC 1620
QY 1621 CTGGCTGCCAACCTTGAAGTCTTCAACTCTCTTCAAGATGAAGATGTCCGTCTATCTCTG 1680
Db 1621 CTGGCTGCCAACCTTGAAGTCTTCAACTCTCTTCAAGATGAAGATGTCCGTCTATCTCTG 1680
QY 1681 GCGCTGCTGCAATGGCTTTGGGGTGTCTCGGAGTCTTCAACACAGTGCACCTTTGGC 1740
Db 1681 GCGCTGCTGCAATGGCTTTGGGGTGTCTCGGAGTCTTCAACACAGTGCACCTTTGGC 1740
QY 1741 CAGAGGCAACCGCTGCTGAGACGCTGCGGAGCTGCGGAGCTCACCTTCTGCTGGGACTCTTC 1800
Db 1741 CAGAGGCAACCGCTGCTGAGACGCTGCGGAGCTGCGGAGCTCACCTTCTGCTGGGACTCTTC 1800
QY 1801 GGTACTCTGCTGTTCTAGTCTATCAAGTGGCTGTGTGTCTGGGCTGCGAGGGCGCGC 1860
Db 1801 GGTACTCTGCTGTTCTAGTCTATCAAGTGGCTGTGTGTCTGGGCTGCGAGGGCGCGC 1860
QY 1861 TCGCCAGCATCTCATCACTTCAATCAACATGTTCTCTTCTCCACAGACCCCGAGCAAC 1920
Db 1861 TCGCCAGCATCTCATCACTTCAATCAACATGTTCTCTTCTCCACAGACCCCGAGCAAC 1920
QY 1921 AGGCTGCTTACCCCGGAGAGGTGTGTCAGGCGCACGCTGTGTGTCTGCGCTTTGGCC 1980
Db 1921 AGGCTGCTTACCCCGGAGAGGTGTGTCAGGCGCACGCTGTGTGTCTGCGCTTTGGCC 1980
QY 1981 ATGTTGCCATCTGCTGTTGGCACACCCCTGACCTGCTGCACCGCCACCGCGCGCGC 2040
Db 1981 ATGTTGCCATCTGCTGTTGGCACACCCCTGACCTGCTGCACCGCCACCGCGCGCGC 2040
QY 2041 CTGCGAGAGGCGCGCTGACCGACAGGAGGAAACAGAGGCGGGTGTGTCGACCTGCTCT 2100
Db 2041 CTGCGAGAGGCGCGCTGACCGACAGGAGGAAACAGAGGCGGGTGTGTCGACCTGCTCT 2100
QY 2101 GACGATCTGTGAATGTGGCTGAGCTCCGATGAGGAAAGGAGGCGGGCTGATGATGAA 2160
Db 2101 GACGATCTGTGAATGTGGCTGAGCTCCGATGAGGAAAGGAGGCGGGCTGATGATGAA 2160
QY 2161 GAGGAGGCGGAGCTGCTCCCTCGAGGTGTCTATGCAACAGGCGCATCCACCATCGAG 2220
Db 2161 GAGGAGGCGGAGCTGCTCCCTCGAGGTGTCTATGCAACAGGCGCATCCACCATCGAG 2220
QY 2221 TTTCTGCTGGGCTGCTTCCAAACACCGCTCTTACTCTGCGCTGTGGGCTTGGGCTTGG 2280
Db 2221 TTTCTGCTGGGCTGCTTCCAAACACCGCTCTTACTCTGCGCTGTGGGCTTGGGCTTGG 2280
QY 2281 GCGCAGCGCCAGCTGCTCGAGGTTCTGTGGGCGATGTGTGATGCGCATAGGCGCTGGGCTG 2340
Db 2281 GCGCAGCGCCAGCTGCTCGAGGTTCTGTGGGCGATGTGTGATGCGCATAGGCGCTGGGCTG 2340
|||||

| | | | | | |
|----|------|---|--|--------------------------|------|
| Qy | 213 | GTTTGA | TGTTTGGCGCTGTGAGAGCTGGAGAGCACTTCACTTCTCTGACGAGGAGGT | 272 | |
| Db | 264 | GTTTGA | TGTTTGGCGCTGTGAGGAGCTGGAGAGACCTTCACTTCTCTGACGAGGAGGT | 323 | |
| Qy | 273 | CGCGCGGCTCGGGCTGGTCTGCTGCCCCCGCCAAAGGGGAGGCTGCCGCA | CCCCACCCCG | 332 | |
| Db | 324 | CGCGCGGCTCGGGCTGGTCTGCTGCCCCCGCCAAAGGGGAGGCTGCCGCA | CCCCACCCCG | 383 | |
| Qy | 333 | GGACCTGTCTGCGCATCCAGGAGGAGACGGAGCGCTTGGCC | CAGAGCTGGGGATGTGCG | 392 | |
| Db | 384 | GGACCTGTCTGCGCATCCAGGAGGAGACGGAGCGCTTGGCC | CAGAGCTGGGGATGTGCG | 443 | |
| Qy | 393 | GGGCAACACAGACGCGCTTGC | CGGCGCCAGCTGCA | CAAGTGGGCGCTCAGAGCGAC | 452 |
| Db | 444 | GGGCAACACAGACGCGCTTGC | CGGCGCCAGCTGCA | CAAGTGGGCGCTCAGAGCGAC | 503 |
| Qy | 453 | ACGCCAGGGCCATGAACCTCAGCTGGCAGCGCGCCCA | CACAGATGGGCGCTCAGAGCGAC | 512 | |
| Db | 504 | ACGCCAGGGCCATGAACCTCAGCTGGCAGCGCGCCCA | CACAGATGGGCGCTCAGAGCGAC | 563 | |
| Qy | 513 | GCCCTGTCTCAGCGCCCCCGGGGGCGGCAC | CAGGACTGAGGGGTCAACTTTTGTGGCAGG | 572 | |
| Db | 564 | GCCCTGTCTCAGCGCCCCCGGGGGCGGCAC | CAGGACTGAGGGGTCAACTTTTGTGGCAGG | 623 | |
| Qy | 573 | TGCCGTGGAGCCCCACAAGCCCCCTGCGCTAGAGCGCTTCTG | AGGGCGCTGCCCGCG | 632 | |
| Db | 624 | TGCCGTGGAGCCCCACAAGCCCCCTGCGCTAGAGCGCTTCTG | AGGGCGCTGCCCGCG | 683 | |
| Qy | 633 | CTTCTCTCATTTGCCAGCTTCAGGGAGCTTGGAGCAGCGCTGGAGCA | CCCCGTGACGGCGGA | 692 | |
| Db | 684 | CTTCTCTCATTTGCCAGCTTCAGGGAGCTTGGAGCAGCGCTGGAGCA | CCCCGTGACGGCGGA | 743 | |
| Qy | 693 | GCCAGCCAGTGGATGACTTCTCTCATCTCTCTGCTGGGTGAGCAGATCGGACAGAGAT | 752 | | |
| Db | 744 | GCCAGCCAGTGGATGACTTCTCTCATCTCTCTGCTGGGTGAGCAGATCGGACAGAGAT | 803 | | |
| Qy | 753 | CCGCAAGATCACGGA | CTGCTTCTCACTGCGCAGCTTCTCCCGTTTCTG | CAGCAGGAGGAGGC | 812 |
| Db | 804 | CCGCAAGATCACGGA | CTGCTTCTCACTGCGCAGCTTCTCCCGTTTCTG | CAGCAGGAGGAGGC | 863 |
| Qy | 813 | CCGCTCTGGGGCCCTGCAGCAGCTGCAACACAGCAGAGCCAGAGCTGAGAGGTCCTCGG | 872 | | |
| Db | 864 | CCGCTCTGGGGCCCTGCAGCAGCTGCAACACAGCAGAGCCAGAGCTGAGAGGTCCTCGG | 923 | | |
| Qy | 873 | GGAGACAGAGCGGTTCTCTGAGCAGAGTGTCTAGGCGCGGTGTG | CAGCTGTCGCGCCAGG | 932 | |
| Db | 924 | GGAGACAGAGCGGTTCTCTGAGCAGAGTGTCTAGGCGCGGTGTG | CAGCTGTCGCGCCAGG | 983 | |
| Qy | 933 | GCAGGTGCAGGTCCACAAGATGAAGCGCGTGTACCTGGCCCTGAAC | CAGTGCAGCGTGAG | 992 | |
| Db | 984 | GCAGGTGCAGGTCCACAAGATGAAGCGCGTGTACCTGGCCCTGAAC | CAGTGCAGCGTGAG | 1043 | |
| Qy | 993 | CACACGCA | CAAGTGCCTCATTTGCCGAGGCTTGTGTCTGTG | GGAGACCTTGCCCGCCCT | 1052 |
| Db | 1044 | CACACGCA | CAAGTGCCTCATTTGCCGAGGCTTGTGTCTGTG | GGAGACCTTGCCCGCCCT | 1103 |
| Qy | 1053 | GCAGGAGCCCTGCGGACAGCTCGATGTGAGGAGGAGTGA | GTGCGTGTCAACCGCAT | 1112 | |
| Db | 1104 | GCAGGAGCCCTGCGGACAGCTCGATGTGAGGAGGAGTGA | GTGCGTGTCAACCGCAT | 1163 | |
| Qy | 1113 | CCCCGCGGAGACATGCC | CCCCCACATCATCCGCA | CAACCGCTTACGCGCCAGCTTCCA | 1172 |
| Db | 1164 | CCCCGCGGAGACATGCC | CCCCCACATCATCCGCA | CAACCGCTTACGCGCCAGCTTCCA | 1223 |
| Qy | 1173 | GGGATCGTGGATCGCTACGGCGTGGCGCGTAC | CCAGGAGGTCA | CCCCCGCTCCCTACAC | 1232 |
| Db | 1224 | GGGATCGTGGATCGCTACGGCGTGGCGCGTAC | CCAGGAGGTCA | CCCCCGCTCCCTACAC | 1283 |
| Qy | 1233 | CATCATCACCTTCCCTTCTGTTTGTGTGA | GTTCGGGATGTGGGCA | CGGGCTGCT | 1292 |
| Db | 1284 | CATCATCACCTTCCCTTCTGTTTGTGTGA | GTTCGGGATGTGGGCA | CGGGCTGCT | 1343 |
| Qy | 1293 | CATGTTCTCTTCCCTTGGCCATG | GTCTTGGCGAGAACGAC | CGCGCTGTGAAAGCCGC | 1352 |

| | | | |
|----|------|---|------|
| Db | 1344 | CATGTTCTCTTTCGGCCCTGGCCATGCTGCTTTCGGAGAAACCGAGCCGGCTGTGAAGCGCGC | 1403 |
| Qy | 1353 | GCAGAACGAGATCTGGCAGACTTTCTTCTCAGGGCGCGCTACCTGCTCCTGCTGTATGGCCCT | 1412 |
| Db | 1404 | GCAGAACGAGATCTGGCAGACTTTCTTCTCAGGGCGCGCTACCTGCTCCTGCTGTATGGCCCT | 1463 |
| Qy | 1413 | GTCTTCATCTACACCGGCTTCATCTACAAACGAGTGCTTCAGTCTGGCGCACACGACATCTT | 1472 |
| Db | 1464 | GTCTTCATCTACACCGGCTTCATCTACAAACGAGTGCTTCAGTCTGGCGCACACGACATCTT | 1523 |
| Qy | 1473 | CCGCTCGGCTGGAGTGTGGCCGCCATGCGCAACAGTCTGGCTGGAGTGATGCAATTCCT | 1532 |
| Db | 1524 | CCGCTCGGCTGGAGTGTGGCCGCCATGCGCAACAGTCTGGCTGGAGTGATGCAATTCCT | 1583 |
| Qy | 1533 | GGCCACGACACGATGCTTACCTCGATCCCAACGTCACCGGTGTCTTCTGGGACCCCTA | 1592 |
| Db | 1584 | GGCCACGACACGATGCTTACCTCGATCCCAACGTCACCGGTGTCTTCTGGGACCCCTA | 1643 |
| Qy | 1593 | CCCTTTGGCATCGATCTCTATTTGGAGCTGGCTGCCAACCACTTGAGCTTCTCTCAACTC | 1652 |
| Db | 1644 | CCCTTTGGCATCGATCTCTATTTGGAGCTGGCTGCCAACCACTTGAGCTTCTCTCAACTC | 1703 |
| Qy | 1653 | CTTCAAGATCAAGATGTCGCTCATCTCTGGCGCTGCTGCAATGGCCCTTTTGGGTGGTCTCT | 1712 |
| Db | 1704 | CTTCAAGATCAAGATGTCGCTCATCTCTGGCGCTGCTGCAATGGCCCTTTTGGGTGGTCTCT | 1763 |
| Qy | 1713 | CGGAGTCTTCAACCAAGTGCACCTTTGGCCAGAGGACACGGCTGCTGCTGGAGACGCTGCC | 1772 |
| Db | 1764 | CGGAGTCTTCAACCAAGTGCACCTTTGGCCAGAGGACACGGCTGCTGCTGGAGACGCTGCC | 1823 |
| Qy | 1773 | GGAGCTCACCTTCTGCTGGAGCTTTTGGTTACCTCGTGTTCCTAGTCTATCTACAAAGTG | 1832 |
| Db | 1824 | GGAGCTCACCTTCTGCTGGAGCTTTTGGTTACCTCGTGTTCCTAGTCTATCTACAAAGTG | 1883 |
| Qy | 1833 | GCTGTGTGTCTGGGCTGCCAGGGCCGCGCTCG --- CCACAGATCCTCATCACTTCATCAA | 1889 |
| Db | 1884 | GCTGTGTGTCTGGGCTGCCAGGGCCGCGCTCGGGCCCCAGCATCTCATCCACTTCATCAA | 1943 |
| Qy | 1890 | CATGTTCTCTTCTCCAGACGCCCAAGCAAGGCTGCTCTACCCCCGGGAGAGAGTGCT | 1949 |
| Db | 1944 | CATGTTCTCTTCTCCAGACGCCCAAGCAAGGCTGCTCTACCCCCGGGAGAGAGTGCT | 2003 |
| Qy | 1950 | CCAGGCCACGCTGGTGCTGGCCCTTGGCCATGTGTGCCCATCTCTGCTGTGGCACACC | 2009 |
| Db | 2004 | CCAGGCCACGCTGGTGCTGGCCCTTGGCCATGTGTGCCCATCTCTGCTGTGGCACACC | 2063 |
| Qy | 2010 | CCTGCACTGCTGCAACCGCCACCGCCCGCTGCGGAGGAGGCCCGCTGACCGACAGGA | 2069 |
| Db | 2064 | CCTGCACTGCTGCAACCGCCACCGCCCGCTGCGGAGGAGGCCCGCTGACCGACAGGA | 2123 |
| Qy | 2070 | GGAAAAAAGGCGCGGGTTGCTGCAACCTGCTGACGCACTGTGTGAATGGCTGGAGCTCCGA | 2129 |
| Db | 2124 | GGAAAAAAGGCGCGGGTTGCTGCAACCTGCTGCAACCTGTGTGAATGGCTGGAGCTCCGA | 2183 |
| Qy | 2130 | TGAGGAAAAAGGCAAGGGGCTTGATGATGAAGAGGAGCGGAGCTGCTCCCTCCAGGT | 2189 |
| Db | 2184 | TGAGGAAAAAGGCAAGGGGCTTGATGATGAAGAGGAGCGGAGCTGCTCCCTCCAGGT | 2243 |
| Qy | 2190 | GCTCATGCAACGAGCCATCCACACCATCGAGTTCTGCTGGGCTGCTCTCCAACACCGC | 2249 |
| Db | 2244 | GCTCATGCAACGAGCCATCCACACCATCGAGTTCTGCTGGGCTGCTCTCCAACACCGC | 2303 |
| Qy | 2250 | CTCCTACCTGGGCTGTGGGCTGTGAGCTTGGCCACGCCCCAGCTGTCCGAGGTTCTGTG | 2309 |
| Db | 2304 | CTCCTACCTGGGCTGTGGGCTGTGAGCTTGGCCACGCCCCAGCTGTCCGAGGTTCTGTG | 2363 |
| Qy | 2310 | GGCCATGGTGATCGCATAGGCTGGGCTTGGGCGGGAGGTGGGCGTGGCGGCTGCTGT | 2369 |
| Db | 2364 | GGCCATGGTGATCGCATAGGCTTGGGCTTGGGCGGGAGGTGGGCGTGGCGGCTGCTGT | 2423 |
| Qy | 2370 | GCTGGTCCCCTATCTTTGCGGCTTTGCGGCTGTGACCGTGGCTATCTCTGCTGGTATGGA | 2429 |

Db 2424 GCTGGTCCCATCTTTGCGGCTTTGCGGTGATGACCGTGGCTATCTCTGCTGGTATGGA 2483
QY 2430 GGGACTCTAGCCTTCTGCAAGCCCTGCGGCTGCACTGGGTGGAATTCAGAACAAAGTT 2489
Db 2484 GGGACTCTAGCCTTCTGCAAGCCCTGCGGCTGCACTGGGTGGAATTCAGAACAAAGTT 2543
QY 2490 CTACTCAGGACAGGGCTACAGCTGAGTCCCTTACCTTCCCTGCGGCTGCACTGGGTGGAATTCAGAACAAAGTT 2549
Db 2544 CTACTCAGGACAGGGCTACAGCTGAGTCCCTTACCTTCCCTGCGGCTGCACTGGGTGGAATTCAGAACAAAGTT 2603
QY 2550 CCCACTGCAGTCTCTGCGAGACCTCTCTGAGCTCTGAGGCGAGGAGGAAATAAGAC 2609
Db 2604 CCCACTGCAGTCTCTGCGAGACCTCTCTGAGCTCTGAGGCGAGGAGGAAATAAGAC 2663
QY 2610 GGTCCGCTTGGCAAAAAA 2637
Db 2664 GGTCCGCTTGGCAAAAAA 2691

RESULT 11
LOCUS Q0718250
DEFINITION Sequence 2582 bp DNA linear PAT 03-FEB-2004
ACCESSION Q0718250
VERSION Q0718250.1 GI:42279107
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 4184 06-SEP-2002; /
PB Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1..2582
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 96.6%; Score 2550.2; DB 2; Length 2582;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2568; Conservative 0; Mismatches 8; Indels 3; Gaps 1;
QY 48 CCGGGGACCATGGGCTCCATGTTCCGAGCGAGGAGGTGGCCCTGCTCAGCTCTTTCT 107
Db 4 CCACAGGACCATGGGCTCCATGTTCCGAGCGAGGAGGTGGCCCTGCTCAGCTCTTTCT 63
QY 108 GCCCAGCGGCTGCTTACACCTGCTGAGTCCGCTGGGCGAGCTGGGCTCTGGAGTT 167
Db 64 GCCCAGCGGCTGCTTACACCTGCTGAGTCCGCTGGGCGAGCTGGGCTCTGGAGTT 123
QY 168 CAGAGACCTCAACGCTCTGGTGAAGCCCTTCCAGAGACGCTTTGTTGATGTTGGCG 227
Db 124 CAGAGACCTCAACGCTCTGGTGAAGCCCTTCCAGAGACGCTTTGTTGATGTTGGCG 183
QY 228 CTGTGAGGAGCTGGAGACCTTCCCTTCTGCGAGGAGGTGGGCGGCTGGGCT 287
Db 184 CTGTGAGGAGCTGGAGACCTTCCCTTCTGCGAGGAGGTGGGCGGCTGGGCT 243
QY 288 GGTCTGCTCCCGCCAAAGGGAGGCTGGCGGACCCCGGAGCTGCTGGCGAT 347
Db 244 GGTCTGCTCCCGCCAAAGGGAGGCTGGCGGACCCCGGAGCTGCTGGCGAT 303
QY 348 CCAGGAGGAGACGGAGCGCTGGCCCGAGGAGCTGGCGGATGTGCGGGGCAACAGCAGGC 407
Db 304 CCAGGAGGAGACGGAGCGCTGGCCCGAGGAGCTGGCGGATGTGCGGGGCAACAGCAGGC 363

QY 408 CTTGCGGGCCAGCTGTGCAACAGCTGTGAGCTTCCAGCGCCGCTGTCTACGCGGCGCATGA 467
Db 364 CTTGCGGGCCAGCTGTGCAACAGCTGTGAGCTTCCAGCGCCGCTGTCTACGCGGCGCATGA 423
QY 468 ACCTCAGCTGTGCGAGCGCCGACACAGATGGGGCTCAGAGAGGACGCGCCCTGCTCCAGGC 527
Db 424 ACCTCAGCTGTGCGAGCGCCGACACAGATGGGGCTCAGAGAGGACGCGCCCTGCTCCAGGC 483
QY 528 CCCCGGGGGCCGCAACAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGGTGAGACCCCA 587
Db 484 CCCCGGGGGCCGCAACAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGGTGAGACCCCA 543
QY 588 CAAGGCCCCCTGCTCTAGAGCGCTGTCTGTGAGGGCTGCGGGCTTCTCATTTGCGAG 647
Db 544 CAAGGCCCCCTGCTCTAGAGCGCTGTCTGTGAGGGCTGCGGGCTTCTCATTTGCGAG 603
QY 648 CTTTCAGGAGCTGTGAGCAGCGCTGTGAGCACCCGCTGACGGGCGAGCCAGCAGTGGAT 707
Db 604 CTTTCAGGAGCTGTGAGCAGCGCTGTGAGCACCCGCTGACGGGCGAGCCAGCAGTGGAT 663
QY 708 GACCTTCTCATCTCTTACTGCGGTGAGCAGATCGGA CAGAAGATCCGCAAGATCACGGA 767
Db 664 GACCTTCTCATCTCTTACTGCGGTGAGCAGATCGGA CAGAAGATCCGCAAGATCACGGA 723
QY 768 CTGCTTCCACTGCTCCAGCTCTTCCGCTTCTGTGAGCAGGAGGAGCCCGCTCGGGGCT 827
Db 724 CTGCTTCCACTGCTCCAGCTCTTCCGCTTCTGTGAGCAGGAGGAGCCCGCTCGGGGCT 783
QY 828 GCAGCAGCTGTCAACAGCAGCAGCAGCAGCTGTGAGGAGTCTCGGGGAGACAGAGCGGTT 887
Db 784 GCAGCAGCTGTCAACAGCAGCAGCAGCAGCTGTGAGGAGTCTCGGGGAGACAGAGCGGTT 843
QY 888 CTTGAGCCAGGCTGTAGGCGCGGTGTGAGCTGTGCTGCGCCAGGCGAGGTGAGGTCCA 947
Db 844 CTTGAGCCAGGCTGTAGGCGCGGTGTGAGCTGTGCTGCGCCAGGCGAGGTGAGGTCCA 903
QY 948 CAAGATGAAGCGCGTGTACTGCGCCCTGAAACAGTGTGAGGCTGAGACACAGCAAGTG 1007
Db 904 CAAGATGAAGCGCGTGTACTGCGCCCTGAAACAGTGTGAGGCTGAGACACAGCAAGTG 963
QY 1008 CTTTCATTTGCGGCGCTGTGCTGTGCGAGACCTGCGCGCTGCGAGGAGCGCTGCG 1067
Db 964 CTTTCATTTGCGGCGCTGTGCTGTGCGAGACCTGCGCGCTGCGAGGAGCGCTGCG 1023
QY 1068 GGA CAGCTCGATGGAGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1127
Db 1024 GGA CAGCTCGATGGAGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1083
QY 1128 GCGCCCGCCACTCATCTGCGACACCAACCGCTTACGGGCGAGCTTCCAGGGCATCTGTGATCG 1187
Db 1084 GCGCCCGCCACTCATCTGCGACACCAACCGCTTACGGGCGAGCTTCCAGGGCATCTGTGATGC 1143
QY 1188 CTACGGCGTGGGCGCTTACAGGAGGTCAACCGCGCTTCCATACCATCATCATCTTCCC 1247
Db 1144 CTACGGCGTGGGCGCTTACAGGAGGTCAACCGCGCTTCCATACCATCATCATCTTCCC 1203
QY 1248 CTTCTCTGTTTGTGTGATGTTTGGGGATGTGGGCGACGCGGCTGCTCATGTTCTTTCG 1307
Db 1204 CTTCTCTGTTTGTGTGATGTTTGGGGATGTGGGCGACGCGGCTGCTCATGTTCTTTCG 1263
QY 1308 CTTGGCCATGTGCTTGGGAGAAACCGACCGGCTGTGAAAGCGCGGAGAACAGATCTG 1367
Db 1264 CTTGGCCATGTGCTTGGGAGAAACCGACCGGCTGTGAAAGCGCGGAGAACAGATCTG 1323
QY 1368 GCAGCTTTCTTCAGGGCGGCTTACCTGCTTCTGCTTATGGGCTGCTTCTCATCTACAC 1427
Db 1324 GCAGCTTTCTTCAGGGCGGCTTACCTGCTTCTGCTTATGGGCTGCTTCTCATCTACAC 1383
QY 1428 CGGCTTCTCATCTCAACAGAGTGTTCAGTGTGCGGCGACAGCATCTTCCCTCGGGCTGGAG 1487
Db 1384 CGGCTTCTCATCTCAACAGAGTGTTCAGTGTGCGGCGACAGCATCTTCCCTCGGGCTGGAG 1443
QY 1488 TGTGGCGCCCATGGGCCAAACAGTGTGCTGGTGGAGTGTGATTCCTTCTGGCGCAGCAGAT 1547

IAEAWCSVRDLPALQEARLSSMBEGVSAVARIIPCRDMPPTLIRNRFASFOGIVD
 AVYQRYEVNPAPIYITFPFLFAVFMFGVHLLMFLFALAWLAENRPAVKAAQN
 ETWQTFPFRGRYLLMLGLPSIYGTIVNECFSEATSIIPSGWSVAAMNQSWDAFL
 AQHTMLTIDPNVTGFLPGYPFGIDPISLWLANHLSPLNSFKKHSVILGVVHMAFGV
 VUGFNHVSFGQRHRLLELPELFTLLGLFGLYFLVLYIKWLCVMAAASAPSILI
 HFINMFLSHSPSNRLLYPRQVQATLVLAAMVPILLGTPLLHLHRHRRLR
 PADQENKAGLLDLPDASVNGWSDESKAGGLDEEAEELVPSVLMHQATHITTEPC
 LQCVNTASYLRMLWLSLAHOLSVLWAMVMRIIGLGLREVGVAAVLVPIFAAFV
 MTVAILLWEGLSAFHLALRLHWBFQNFYSGTGYKLSPTFAATDD"

ORIGIN

| Query Match | | 78.3%; | Score 2068; | DB 2; | Length 2488; |
|-----------------------|------|---|--------------|----------------|-------------------|
| Best Local Similarity | | 98.0%; | Pred. No. 0; | | |
| Matches 2105; | | Conservative | 0; | Mismatches 40; | Indels 3; Gaps 1; |
| QY | 496 | GGGGCTCAGAGGACGCCCTGCTCAGAGCCCCCGGGGGCCGACACAGAACCTGAGG | 555 | | |
| DB | 327 | GGGGCCCTGGTGGGGAGGACGGGAGGAGGTGGAGCAGCGCTGCCACGCCCGTGGCC | 386 | | |
| QY | 556 | GTCAACTTTGTGGCAGGTGCGTGGAGCCCAAGAGCCCTGCCCTAGAGCGCTGCTC | 615 | | |
| DB | 387 | GCCAGCTTTGTGCAAGTGGCTGGAGCCCAAGAGCCCTGCCCTAGAGCGCTGCTC | 446 | | |
| QY | 616 | TGAGGGCTGCGCGGGTTCCTCATTTGCCAGCTTCAGGGAGCTGGAGCAGCGCTGGAG | 675 | | |
| DB | 447 | TGAGGGCTGCGCGGGTTCCTCATTTGCCAGCTTCAGGGAGCTGGAGCAGCGCTGGAG | 506 | | |
| QY | 676 | CACCCGCTGACGGGAGCCAGCAGCTGGATGACCTTCTCATCTCTACTTGGGGTGAG | 735 | | |
| DB | 507 | CACCCGCTGACGGGAGCCAGCAGCTGGATGACCTTCTCATCTCTACTTGGGGTGAG | 566 | | |
| QY | 736 | CAGATCGGACAGAAGATCGCAAGATCACGGACTGCTTCCACTGCCAGCTTTCCTGTT | 795 | | |
| DB | 567 | CAGATCGGACAGAAGATCGCAAGATCACGGACTGCTTCCACTGCCAGCTTTCCTGTT | 626 | | |
| QY | 796 | GTGACAGGAGGAGCGCGCTCGGGGCCCTGACAGCAGCTGCAACAGCAGACCGAGG | 855 | | |
| DB | 627 | GTGACAGGAGGAGCGCGCTCGGGGCCCTGACAGCAGCTGCAACAGCAGACCGAGG | 686 | | |
| QY | 856 | CTGACGAGGAGTCTCGGGAGACAGAGGTTCTTGGAGCAGGTGCTAGGCCGGGTGCTG | 915 | | |
| DB | 687 | CTGACGAGGAGTCTCGGGAGACAGAGGTTCTTGGAGCAGGTGCTAGGCCGGGTGCTG | 746 | | |
| QY | 916 | CAGCTGCTGCCGCCAGGGCAGGTGACAGTCCCAAGAGTGAAGCCGTGTACTTGGCCCTG | 975 | | |
| DB | 747 | CAGCTGCTGCCGCCAGGGCAGGTGACAGTCCCAAGAGTGAAGCCGTGTACTTGGCCCTG | 806 | | |
| QY | 976 | AACCAAGTGAAGGTGAGCACACGCAAGTGCCCTCATTTGCCAGGCTTGGTCTGTG | 1035 | | |
| DB | 807 | AACCAAGTGAAGGTGAGCACACGCAAGTGCCCTCATTTGCCAGGCTTGGTCTGTG | 866 | | |
| QY | 1036 | CGAGACCTGCCGCCCTCGAGAGGCCCTCGGGACAGCTCGATGGAGGAGGTGAGT | 1095 | | |
| DB | 867 | CGAGACCTGCCGCCCTCGAGAGGCCCTCGGGACAGCTCGATGGAGGAGGTGAGT | 926 | | |
| QY | 1096 | GCCGTGGCTCACCGCATCCCTGCCGGACATGCCGCCACACTCATCCGACCAACCGC | 1155 | | |
| DB | 927 | GCCGTGGCTCACCGCATCCCTGCCGGACATGCCGCCACACTCATCCGACCAACCGC | 986 | | |
| QY | 1156 | TTACAGGCGAGTTTCCAGGGCATTCGTGATTCGCTACGGGGTGGCGCTTACAGAGGTC | 1215 | | |
| DB | 987 | TTACAGGCGAGTTTCCAGGGCATTCGTGATTCGCTACGGGGTGGCGCTTACAGAGGTC | 1046 | | |
| QY | 1216 | AACCCGCTCCCTACACCATCATACCTTCCCTTCCCTTCTGCTGATGTTTCGGGAT | 1275 | | |
| DB | 1047 | AACCCGCTCCCTACACCATCATACCTTCCCTTCCCTTCTGCTGATGTTTCGGGAT | 1106 | | |
| QY | 1276 | GTGGGCCACGGGCTGCTCATGTTCTCTTTCGCCCTGGGCATGGTCTTTCGGAGAACCGA | 1335 | | |
| DB | 1107 | GTGGGCCACGGGCTGCTCATGTTCTCTTTCGCCCTGGGCATGGTCTTTCGGAGAACCGA | 1166 | | |
| QY | 1336 | CGGGCTGTGAAGCCGCGAGAACAGATCTGGCAGACTTCTTTCAGGGGCCGCTACCTG | 1395 | | |

| | | | | | |
|----|------|--|------|--|--|
| DB | 1167 | COGCTGTGAAGGGCGCGCAGAAACAGATCTGGCAGACTTTTCTTTCAGGGGCCGTACCTG | 1226 | | |
| QY | 1396 | CTCCTGCTTATGGGGCTGTTCTTCAATCTACACCGGCTTCACTACAAACGAGTGTCTTCA | 1455 | | |
| DB | 1227 | CTCCTGCTTATGGGGCTGTTCTTCAATCTACACCGGCTTCACTACAAACGAGTGTCTTCA | 1286 | | |
| QY | 1456 | CGCGCCACACGACTCTTCCCTCGGGCTGGAGTGTGGCGCCATCGCCAAACAGTCTGGC | 1515 | | |
| DB | 1287 | CGCGCCACACGACTCTTCCCTCGGGCTGGAGTGTGGCGCCATCGCCAAACAGTCTGGC | 1346 | | |
| QY | 1516 | TGAGTGTGATTCCTTGGCCCCAGCACAGATGCTTACCTTGGATCCCAACGCTCACCGT | 1575 | | |
| DB | 1347 | TGAGTGTGATTCCTTGGCCCCAGCACAGATGCTTACCTTGGATCCCAACGCTCACCGT | 1406 | | |
| QY | 1576 | GTCTTCTTGGGACCTTACCCCTTTTGGCATCGATCTCTATTTGGAGCTTGGCTTGCACAC | 1635 | | |
| DB | 1407 | GTCTTCTTGGGACCTTACCCCTTTTGGCATCGATCTCTATTTGGAGCTTGGCTTGCACAC | 1466 | | |
| QY | 1636 | TTGAGCTTCTCAACTCTCTTCAAGATGAAGATGTCGCTCATCTGGGGCTGCTGCACATG | 1695 | | |
| DB | 1467 | TTGAGCTTCTCAACTCTCTTCAAGATGAAGATGTCGCTCATCTGGGGCTGCTGCACATG | 1526 | | |
| QY | 1696 | GCCTTGGGGTGTCTCGAGTCTTCAACCAACGTCATCTTTGGCCAGAGCACCGGCTG | 1755 | | |
| DB | 1527 | GCCTTGGGGTGTCTCGAGTCTTCAACCAACGTCATCTTTGGCCAGAGCACCGGCTG | 1586 | | |
| QY | 1756 | CTGCTGGAGACGCTGCCGAGCTCACCTTCTGCTGGGACTCTTTCGGTTCCTGCTGTTTC | 1815 | | |
| DB | 1587 | CTGCTGGAGACGCTGCCGAGCTCACCTTCTGCTGGGACTCTTTCGGTTCCTGCTGTTTC | 1646 | | |
| QY | 1816 | CTAGTCATCTCAAGTGGCTGTGCTGCGGTGCGCAGGGCGGCTCG -- -CCACAGCATC | 1872 | | |
| DB | 1647 | CTAGTCATCTCAAGTGGCTGTGCTGCGGTGCGCAGGGCGGCTCGGCCCCCAGCATC | 1706 | | |
| QY | 1873 | CTCATCCACTTCAATCAACATGTTCTTCTTCCACAGCCCCAGCAACAGGCTGCTCTAC | 1932 | | |
| DB | 1707 | CTCATCCACTTCAATCAACATGTTCTTCTTCCACAGCCCCAGCAACAGGCTGCTCTAC | 1766 | | |
| QY | 1933 | CCCCGGCAGGAGTGGTTCAGGCCACGCTGCTGCTGCTGGGCTTGGCCATGCTGCCCCATC | 1992 | | |
| DB | 1767 | CCCCGGCAGGAGTGGTTCAGGCCACGCTGCTGCTGCTGGGCTTGGCCATGCTGCCCCATC | 1826 | | |
| QY | 1993 | CTGCTGCTTGGCACACCCCTGCACTGCTGCAACCGCCACCGCGCGCTGCGGAGGAGG | 2052 | | |
| DB | 1827 | CTGCTGCTTGGCACACCCCTGCACTGCTGCAACCGCCACCGCGCGCTGCGGAGGAGG | 1886 | | |
| QY | 2053 | CCGCTGACCGACAGGAGGAAAAAAGAGCCCGGGTTGTGGACCTTGCCTGACGATCTGTG | 2112 | | |
| DB | 1887 | CCGCTGACCGACAGGAGGAAAAAAGAGCCCGGGTTGTGGACCTTGCCTGACGATCTGTG | 1946 | | |
| QY | 2113 | AATGGCTGGAGTCCGATGAGGAAAGCAGGGGGCTGGATGATGAGAGGAGGCGCGAG | 2172 | | |
| DB | 1947 | AATGGCTGGAGTCCGATGAGGAAAGCAGGGGGCTGGATGATGAGAGGAGGCGCGAG | 2006 | | |
| QY | 2173 | CTGCTGCCCTCCGAGGTGCTCATGACACAGGCGCATCCACACATCGAGTCTTGCCTGGGC | 2232 | | |
| DB | 2007 | CTGCTGCCCTCCGAGGTGCTCATGACACAGGCGCATCCACACATCGAGTCTTGCCTGGGC | 2066 | | |
| QY | 2233 | TGCGTCTCCAAACACCGCTCTCTACTGCGCTGTGGGGCTTGGAGCTTGGCCACCGCCAG | 2292 | | |
| DB | 2067 | TGCGTCTCCAAACACCGCTCTCTACTGCGCTGTGGGGCTTGGAGCTTGGCCACCGCCAG | 2126 | | |
| QY | 2293 | CTGCTCGAGGTTCGTGGGCCATGCTGATGCGCATAGGCTTGGGCTGGGCGGGAGGTG | 2352 | | |
| DB | 2127 | CTGCTCGAGGTTCGTGGGCCATGCTGATGCGCATAGGCTTGGGCTGGGCGGGAGGTG | 2186 | | |
| QY | 2353 | GCGCTGGCGGCTGCTGCTGCTGCCATCTTTTTCGCGCTTTTTCGCTGATGACCGTGGCT | 2412 | | |
| DB | 2187 | GCGCTGGCGGCTGCTGCTGCTGCCATCTTTTTCGCGCTTTTTCGCTGATGACCGTGGCT | 2246 | | |
| QY | 2413 | ATCTGCTGTGTGATGAGGAGGACTCTCAGCTTCTTGCACGCGCTTGGGCTGACCTGGGTG | 2472 | | |
| DB | 2247 | ATCTGCTGTGTGATGAGGAGGACTCTCAGCTTCTTGCACGCGCTTGGGCTGACCTGGGTG | 2306 | | |

507 CACCCCGTGAACGGCGAGCCAGCCACGTGGATGACCTTCTCTCATCTCTACTTGGGGTGAG 566
736 CAGATCGGACAGAAAGATCCCAAGATACGGAGCTGTTCCACTGCGCAAGTCTTCCCGTTT 795
567 CAGATCGGACAGAAAGATCCCAAGATACCGAGCTGTTCCACTGCGCAAGTCTTCCCGTTT 626
796 CTGCAGCAGGAGGAGGCGCGCTCGGGGCCCTGCAGCAGTGCACAGCAGAGCCAGAG 855
627 CTGCAGCAGGAGGAGGCGCGCTCGGGGCCCTGCAGCAGTGCACAGCAGAGCCAGGAG 686
856 CTGCAGGAGGTCCTCGGGGACAGAGCGGTTCTGAGCCAGAGTGTAGCCCGGGTGTCTG 915
687 CTGCAGGAGGTCCTCGGGGAGACAGAGCGGTTCTGAGCCAGAGTGTAGCCCGGGTGTCTG 746
916 CAGCTCTGCCGCCAGGGCAGGTGCAGGTCCACAAGATGAAGGCCGTGTACCTTGGCCCTG 975
747 CAGCTCTGCCGCCAGGGCAGGTGCAGGTCCACAAGATGAAGGCCGTGTACCTTGGCCCTG 806
976 AACCACTGACAGCGTGAGCACCAAGTGCCTCATTTGCCGAGGCGTGGTGTCTGTG 1035
807 AACCACTGACAGCGTGAGCACCAAGTGCCTCATTTGCCGAGGCGTGGTGTCTGTG 866
1036 CGAGACCTGCCGCCCTGACAGGAGCCCTGCGGACAGTGCATGAGGAGGAGGAGTGA 1095
867 CGAGACCTGCCGCCCTGACAGGAGCCCTGCGGACAGTGCATGAGGAGGAGGAGTGA 926
1096 GCGGTGCTCACCGCATCCCTCTGCCGGGACATGCCCCACACTCATCCGACCAACCGC 1155
927 GCGGTGCTCACCGCATCCCTCTGCCGGGACATGCCCCACACTCATCCGACCAACCGC 986
1156 TTACCGGCGAGCTTCGAGGCGATCTGGATCGCTACGCGGTGGGCGCTTACAGAGATC 1215
987 TTACCGGCGAGCTTCGAGGCGATCTGGATCGCTACGCGGTGGGCGCTTACAGAGATC 1046
1216 AACCCTGCTCCCTACACCATCATCATCTGCCCTTCCTGTTGCTGTGATGTTGCGGAT 1275
1047 AACCCTGCTCCCTACACCATCATCATCTGCCCTTCCTGTTGCTGTGATGTTGCGGAT 1106
1276 GTGGGCGAGGGCTGCTCATGTTCTCTTGCCCTTGCCCATGTCCTTTCAGGGGCGCTACCTG 1395
1107 GTGGGCGAGGGCTGCTCATGTTCTCTTGCCCTTGCCCATGTCCTTTCAGGGGCGCTACCTG 1226
1336 CCGGCTGTCAAGCGCGCGAGAACAGATCTGGCAGACTTCTTCAGGGGCGCTACCTG 1455
1167 CCGGCTGTGAAGCGCGCGAGAACAGATCTGGCAGACTTCTTCAGGGGCGCTACCTG 1286
1396 CTCTCGTTATGGGCTGTCTTCCATCTACACCGGCTTCATCTACACGAGTGTCTCAGT 1396
1327 CTCTCGTTATGGGCTGTCTTCCATCTACACCGGCTTCATCTACACGAGTGTCTCAGT 1286
1456 CGGCGCACAGCATCTTCCCTCGGCTGGAGTGTGGCGCCATGGCCAAACAGTCTGGC 1515
1287 CGGCGCACAGCATCTTCCCTCGGCTGGAGTGTGGCGCCATGGCCAAACAGTCTGGC 1346
1516 TGGAGTGATGATTCCTTGGGCCAGACACAGTGTTCACCTGGATCCCAAGTCAACCGGT 1575
1347 TGGAGTGATGATTCCTTGGGCCAGACACAGTGTTCACCTGGATCCCAAGTCAACCGGT 1406
1576 GTCTTCTGGGACCTTACCTTGGCATCGATCTTATTTGGAGCGTGGCTGCCAACAC 1635
1407 GTCTTCTGGGACCTTACCTTGGCATCGATCTTATTTGGAGCGTGGCTGCCAACAC 1466
1636 TTGAGCTTCTCAACTCTTCAAGATGAAGTCCGTGCATCTCGGCGCTCGTGCACATG 1695
1467 TTGAGCTTCTCAACTCTTCAAGATGAAGTCCGTGCATCTCGGCGCTCGTGCACATG 1526
1696 GCGTTTGGGGTGTCTCGGAGTCTTCAACCAAGTGAATTTGGCCAGAGGCACCGGCTG 1755
1527 GCGTTTGGGGTGTCTCGGAGTCTTCAACCAAGTGAATTTGGCCAGAGGCACCGGCTG 1586
1756 CTGCTGGAGACGCTGCGGAGGCTCATCTTCTGCTGGGACTCTTTCGGTGTACCTCGTGTTC 1815

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2006, 03:45:01 ; Search time 1504 Seconds
(without alignments)
12238.521 Million cell updates/sec

Title: US-10-783-519-1

Perfect score: 2640

Sequence: 1 cggcgtcgccgagcggcag.....gcaaaaaaaaaaaaaaa 2640

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*
15: Geneseqn2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 2640 | 100.0 | 2640 | 2 | AAV04529 Human OC- |
| 2 | 2640 | 100.0 | 2640 | 6 | ABN86735 Human OC- |
| 3 | 2640 | 100.0 | 2640 | 13 | ADR44118 Human 116 |
| 4 | 2640 | 100.0 | 2640 | 14 | ADY92605 Human 116 |
| 5 | 2640 | 100.0 | 2655 | 6 | ABL64400 Stomach c |
| 6 | 2640 | 100.0 | 2655 | 6 | ABL68345 Kidney ca |
| 7 | 2640 | 100.0 | 2655 | 6 | ABK84337 Human cdn |
| 8 | 2640 | 100.0 | 2655 | 6 | ABN96866 Gene #336 |
| 9 | 2640 | 100.0 | 2655 | 10 | ACA56496 Human sig |
| 10 | 2640 | 100.0 | 2655 | 11 | ADI31590 Human cdn |
| 11 | 2640 | 100.0 | 2655 | 12 | ADI56292 Human pol |
| 12 | 2640 | 100.0 | 2655 | 13 | ADS83657 Human lym |
| 13 | 2592.2 | 98.2 | 2700 | 12 | ADI28819 Human mod |
| 14 | 2592.2 | 98.2 | 2700 | 12 | ADP10401 Reference |
| 15 | 2592.2 | 98.2 | 2700 | 13 | ACN39267 Tumour-as |
| 16 | 2577.4 | 97.6 | 2676 | 13 | ADQ86432 Human tum |
| 17 | 2577.4 | 97.6 | 2676 | 13 | ADQ83344 Human tum |
| 18 | 2575.8 | 97.6 | 2676 | 12 | ADQ85289 Human tum |

| | | | | | |
|----|--------|------|------|----|-----------|
| 19 | 2470.6 | 93.6 | 2493 | 14 | ADV43978 |
| 20 | 2327.6 | 88.2 | 2561 | 12 | ADR22582 |
| 21 | 2310.4 | 87.5 | 2762 | 13 | ACN43472 |
| 22 | 2068 | 78.3 | 2488 | 2 | AAAX24912 |
| 23 | 2066.4 | 78.3 | 2488 | 2 | AAAX24913 |
| 24 | 2063.2 | 78.2 | 2480 | 13 | ACN39268 |
| 25 | 1989.4 | 75.4 | 3104 | 6 | ABQ54601 |
| 26 | 1824.2 | 69.1 | 1845 | 14 | ADV43979 |
| 27 | 742.8 | 28.1 | 2937 | 14 | ADZ62313 |
| 28 | 741.2 | 28.1 | 2945 | 2 | AAQ38591 |
| 29 | 637.4 | 24.1 | 3876 | 10 | ADP72690 |
| 30 | 637.4 | 24.1 | 3876 | 12 | ADP72690 |
| 31 | 637.4 | 24.1 | 3876 | 13 | ADV41430 |
| 32 | 635.8 | 24.1 | 4607 | 5 | ABV21479 |
| 33 | 635.8 | 24.1 | 4607 | 5 | ABV27297 |
| 34 | 634.2 | 24.0 | 2718 | 12 | ADI28820 |
| 35 | 634.2 | 24.0 | 2718 | 13 | ADR25554 |
| 36 | 634.2 | 24.0 | 2718 | 13 | ADP54580 |
| 37 | 634.2 | 24.0 | 2718 | 13 | ADP23896 |
| 38 | 634.2 | 24.0 | 2718 | 14 | ADY19961 |
| 39 | 634.2 | 24.0 | 2718 | 14 | ADY15609 |
| 40 | 618.6 | 23.4 | 3062 | 6 | ABZ12013 |
| 41 | 618.6 | 23.4 | 3062 | 12 | ADM44531 |
| 42 | 617.2 | 23.4 | 3137 | 12 | ADI28821 |
| 43 | 617.2 | 23.4 | 3137 | 14 | AEC83090 |
| 44 | 612 | 23.2 | 612 | 10 | ADD27224 |
| 45 | 609.4 | 23.1 | 3014 | 4 | AAF27720 |

ALIGNMENTS

RESULT 1

AAV04529 standard; cDNA; 2640 BP.

AC AAV04529;

DT 02-JUL-1998 (first entry)

DE Human OC-116 kDa cDNA sequence.

KW Osteoclast; human; OC-116 kDa; screening; cell surface marker; probe; ss.

OS Homo sapiens.

PH Key Location/Qualifiers

FT CDS 58..2523

FT /*tag= a

FT /product= "OC-116 kDa"

FT /note= "no stop codon given"

XX WO9803651-A1.

XX 29-JAN-1998.

XX 10-JUL-1997; 97WO-US012569.

XX 19-JUL-1996; 96US-00684932.

XX (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.

XX Stashenko P, Li Y, Wucherpennig AL;

XX WPI; 1998-120776/11.

XX P-PSDB; AAW41943.

XX Osteoclast specific or related DNA sequence - useful as probe to screen

XX genomic DNA or cDNA library, or as osteoclast cell surface marker.

XX Example 8; Page 58-61; 75pp; English.

XX The present sequence represents the human OC-116 kDa cDNA sequence which

CC is used in an example of the present invention, which describes isolated
CC osteoclast specific or related DNA sequences. The present invention also
CC describes; a DNA construct capable of replicating and optionally
CC expressing, in a host cell, osteoclast specific or related DNA,
CC comprising an osteoclast specific or related DNA sequence and sequences
CC necessary for transforming or transfecting a host cell, and for
CC replicating and optionally expressing an osteoclast specific or related
CC DNA sequence in a host cell; and a cell stably transformed or transfected
CC with the DNA construct. The osteoclast specific or related DNA sequence
CC can be used as a probe to screen a genomic DNA or cDNA library for
CC osteoclast specific or related DNA sequences, or as an osteoclast cell
CC surface marker
XX

8Q Sequence 2640 BP; 445 A; 885 C; 824 G; 486 T; 0 U; 0 Other;

| | Query Match | 100.0%; | Score 2640; | DB 2; | Length 2640; |
|----|-----------------------|--|---------------|-----------|--------------|
| | Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| | Matches 2640; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| QY | 1 | CGGCGTGGCGGACGGGACGAGCGAGCGGCGGCGGAGCACACCCGGGGACCATG | 60 | | |
| DB | 1 | CGGCGTGGCGGACGGGACGAGCGGCGGCGGAGCACACCCGGGGACCATG | 60 | | |
| QY | 61 | GGCTCCATGTTCCGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCACAGCGGT | 120 | | |
| DB | 61 | GGCTCCATGTTCCGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCACAGCGGT | 120 | | |
| QY | 121 | GGCTACACCTGGTGGTGGGCGGAGCTGGGCTCGTGGAGTTGAGAGCTCAAC | 180 | | |
| DB | 121 | GGCTACACCTGGTGGTGGGCGGAGCTGGGCTCGTGGAGTTGAGAGCTCAAC | 180 | | |
| QY | 181 | GGCTCGGTGAGGCGCTTCAGAGACGCTTTGTGTGTGATGTTGGGCTGTGAGGAGCTG | 240 | | |
| DB | 181 | GGCTCGGTGAGGCGCTTCAGAGACGCTTTGTGTGTGATGTTGGGCTGTGAGGAGCTG | 240 | | |
| QY | 241 | GAGAGACCTTCACTTCTGAGGAGGAGGTGCGCGGGCTGGGTGCTCTGCCGCCCG | 300 | | |
| DB | 241 | GAGAGACCTTCACTTCTGAGGAGGAGGTGCGCGGGCTGGGTGCTCTGCCGCCCG | 300 | | |
| QY | 301 | CCAAAGGGAGGCTGCGGACACCCACCCCGGACCTGCTGGCATCCAGGAGGAGCG | 360 | | |
| DB | 301 | CCAAAGGGAGGCTGCGGACACCCACCCCGGACCTGCTGGCATCCAGGAGGAGCG | 360 | | |
| QY | 361 | GAGCGCTGGCGGAGCTGGGGATGTGCGGGCAACAGCAGGCGCTGCGGGGCCAG | 420 | | |
| DB | 361 | GAGCGCTGGCGGAGCTGGGGATGTGCGGGCAACAGCAGGCGCTGCGGGGCCAG | 420 | | |
| QY | 421 | CTGCAACAGCTGAGCTCCA CGCGCGGTGCTACGCGAGGCGCATGAACTCAGCTGGCA | 480 | | |
| DB | 421 | CTGCAACAGCTGAGCTCCA CGCGCGGTGCTACGCGAGGCGCATGAACTCAGCTGGCA | 480 | | |
| QY | 481 | GCGGCCACACAGATGGGSCCTCAGAGAGCGCCCTGCTCCAGGCCCGGGGGCGG | 540 | | |
| DB | 481 | GCGGCCACACAGATGGGSCCTCAGAGAGCGCCCTGCTCCAGGCCCGGGGGCGG | 540 | | |
| QY | 541 | CACGAGGACTGAGGGTCAACTTTGTGGCAGGTGCGGTGGAGCCCGCACAGGCGCTGCC | 600 | | |
| DB | 541 | CACGAGGACTGAGGGTCAACTTTGTGGCAGGTGCGGTGGAGCCCGCACAGGCGCTGCC | 600 | | |
| QY | 601 | CTAGAGCGCTGCTCTGGAGGGCTGCGCGGCTTCTCATTTGCCAGCTTCAGGAGCTG | 660 | | |
| DB | 601 | CTAGAGCGCTGCTCTGGAGGGCTGCGCGGCTTCTCATTTGCCAGCTTCAGGAGCTG | 660 | | |
| QY | 661 | GAGCAGCGCTGGAGCACCCCTGACGGGAGCGGACCGACCTGGATGACCTTCTCATC | 720 | | |
| DB | 661 | GAGCAGCGCTGGAGCACCCCTGACGGGAGCGGACCGACCTGGATGACCTTCTCATC | 720 | | |
| QY | 721 | TCCTACTGGGGTGAGCAGATCGGACAGAGAGTCCGCAAGATCACGGACTGCTTCCACTGC | 780 | | |
| DB | 721 | TCCTACTGGGGTGAGCAGATCGGACAGAGAGTCCGCAAGATCACGGACTGCTTCCACTGC | 780 | | |
| QY | 781 | CAGCTTCTCCGTTCTGAGGAGGAGGCGGCTCGGGGCGCTGAGGAGCTGCAA | 840 | | |

| | | | |
|----|------|--|------|
| DB | 781 | CAGCTTCTCCGTTTCTGCAGCAGGAGGAGGCGGCTCGGGGCGCTGCGAGCAGCTGCAA | 840 |
| QY | 841 | CAGCAGAGCCAGGAGCTGCAGAGGCTCTCGGGGAGACAGAGCGGTTCTTGAGCCAGGTG | 900 |
| DB | 841 | CAGCAGAGCCAGGAGCTGCAGAGGCTCTCGGGGAGACAGAGCGGTTCTTGAGCCAGGTG | 900 |
| QY | 901 | CTAGGCCGGGTGCTCGAGCTGCTCGCGCCAGGCGAGGTGCAGGTCCACAAGATGAAGGCC | 960 |
| DB | 901 | CTAGGCCGGGTGCTCGAGCTGCTCGCGCCAGGCGAGGTGCAGGTCCACAAGATGAAGGCC | 960 |
| QY | 961 | GTGTACTGGCCCTGAACAGTGCAGCTGAGCACCAAGAGTGCCTCAATGCGGAG | 1020 |
| DB | 961 | GTGTACTGGCCCTGAACAGTGCAGCTGAGCACCAAGAGTGCCTCAATGCGGAG | 1020 |
| QY | 1021 | GCTCGGTGCTCTGTGCGAGACCTGCGGCCCTGCGAGGAGGCGCTGCGGAGCAGCTCGATG | 1080 |
| DB | 1021 | GCTCGGTGCTCTGTGCGAGACCTGCGGCCCTGCGAGGAGGCGCTGCGGAGCAGCTCGATG | 1080 |
| QY | 1081 | GAGGAGGAGGTGAGTGGCGGTCAACCGCATCCCTGCGGGGACATGCCCCCACCACATC | 1140 |
| DB | 1081 | GAGGAGGAGGTGAGTGGCGGTCAACCGCATCCCTGCGGGGACATGCCCCCACCACATC | 1140 |
| QY | 1141 | ATCCGACCAACCCGCTTCAOGGCCAGTTCAGAGGCGATCGTGGATCGCTACGCGGTGGGC | 1200 |
| DB | 1141 | ATCCGACCAACCCGCTTCAOGGCCAGTTCAGAGGCGATCGTGGATCGCTACGCGGTGGGC | 1200 |
| QY | 1201 | CGCTACCAAGGAGTCAACCCCGCTCCCTACACCATCATCACCTTCCCTTCTGTTGCT | 1260 |
| DB | 1201 | CGCTACCAAGGAGTCAACCCCGCTCCCTACACCATCATCACCTTCCCTTCTGTTGCT | 1260 |
| QY | 1261 | GTGATGTTTCGGGGATGTGGGCCACCGGCTGCTCATGTTCTCTTGGCCCTGGCCATGGTC | 1320 |
| DB | 1261 | GTGATGTTTCGGGGATGTGGGCCACCGGCTGCTCATGTTCTCTTGGCCCTGGCCATGGTC | 1320 |
| QY | 1321 | CTTTCGGGAGAACCGACCGGCTGTGAAAGCGCGGAGAGAGATCTGGCAGACTTCTTTC | 1380 |
| DB | 1321 | CTTTCGGGAGAACCGACCGGCTGTGAAAGCGCGGAGAGAGATCTGGCAGACTTCTTTC | 1380 |
| QY | 1381 | AGGGGCGCTACCTGCTCTGCTTATGGGCTGTTCTCCATCTACACGGCTTCATCTAC | 1440 |
| DB | 1381 | AGGGGCGCTACCTGCTCTGCTTATGGGCTGTTCTCCATCTACACGGCTTCATCTAC | 1440 |
| QY | 1441 | AACGAGTGTCTCAGTCGCGCCACACGAGCATCTTCCCTCGGGCTGGAGTGTGGCGGCATG | 1500 |
| DB | 1441 | AACGAGTGTCTCAGTCGCGCCACACGAGCATCTTCCCTCGGGCTGGAGTGTGGCGGCATG | 1500 |
| QY | 1501 | GCACCAAGTCTGCTGAGTGTGATTCCTGCGGCCAGCACAGATGCTTACCTTGAT | 1560 |
| DB | 1501 | GCACCAAGTCTGCTGAGTGTGATTCCTGCGGCCAGCACAGATGCTTACCTTGAT | 1560 |
| QY | 1561 | CCCAAGCTCACCGGTGCTCTCTGGGAGCCCTACCCCTTGGGATCGATCTTATTTGGAGC | 1620 |
| DB | 1561 | CCCAAGCTCACCGGTGCTCTCTGGGAGCCCTACCCCTTGGGATCGATCTTATTTGGAGC | 1620 |
| QY | 1621 | CTGGCTGCCAACCACTTGAGTCTTCAACTCTTCAAGATGAAGATGTCCTATTCCTG | 1680 |
| DB | 1621 | CTGGCTGCCAACCACTTGAGTCTTCAACTCTTCAAGATGAAGATGTCCTATTCCTG | 1680 |
| QY | 1681 | GGCTGCTGCAATGGCCTTTGGGGTGTGCTCGAGTCTTCAACCAAGTGCATTTGGC | 1740 |
| DB | 1681 | GGCTGCTGCAATGGCCTTTGGGGTGTGCTCGAGTCTTCAACCAAGTGCATTTGGC | 1740 |
| QY | 1741 | CAGAGGCAACCGCTGCTGAGACGCTGCGGAGCTCACCTTCTGCTGGGACTCTTC | 1800 |
| DB | 1741 | CAGAGGCAACCGCTGCTGAGACGCTGCGGAGCTCACCTTCTGCTGGGACTCTTC | 1800 |
| QY | 1801 | GGTTACCTCGTGTCTCTAGTCACTTACAGTGGCTGTGTCTGGGCTGCGGAGGCGGCC | 1860 |
| DB | 1801 | GGTTACCTCGTGTCTCTAGTCACTTACAGTGGCTGTGTCTGGGCTGCGGAGGCGGCC | 1860 |
| QY | 1861 | TGCCCCAGGATCTCATCACTTCAATCAAGTGTCTTCTTCTCCACAGCCCCAGCAAC | 1920 |
| DB | 1861 | TGCCCCAGGATCTCATCACTTCAATCAAGTGTCTTCTTCTCCACAGCCCCAGCAAC | 1920 |

```
QY 1921 AGGCTGCTTACCCCGGCGAGGAGTGGTCCAGGCGACGCTGGTGGTCTTGGGCTTGGCC 1980
DB 1921 AGGCTGCTTACCCCGGCGAGGAGTGGTCCAGGCGACGCTGGTGGTCTTGGGCTTGGCC 1980
QY 1981 ATGGTCCCATCTGCTGCTTGGCACACCCCTGCACCTGCTGACCGCCACCGCCGCGC 2040
DB 1981 ATGGTCCCATCTGCTGCTTGGCACACCCCTGCACCTGCTGACCGCCACCGCCGCGC 2040
QY 2041 CTGCGGAGGAGGCGGCTGACCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100
DB 2041 CTGCGGAGGAGGCGGCTGACCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100
QY 2101 GAGGCATCTGTGATGCTGGAGCTCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
DB 2101 GAGGCATCTGTGATGCTGGAGCTCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
QY 2161 GAGGAGGCGGAGCTGCTGCTCCCGAGGCTCATGACGAGGAGGAGGAGGAGGAGGAGG 2220
DB 2161 GAGGAGGCGGAGCTGCTGCTCCCGAGGCTCATGACGAGGAGGAGGAGGAGGAGGAGG 2220
QY 2221 TTGCTGCTGGGCTGCTTCCACACCGGCTTCTACCTGCGGCTGTGGGCTTGGGCTTGG 2280
DB 2221 TTGCTGCTGGGCTGCTTCCACACCGGCTTCTACCTGCGGCTGTGGGCTTGGGCTTGG 2280
QY 2281 GCCACGCGGAGCTGCTGAGGCTTCTGAGGCTTCTGAGGCTTCTGAGGCTTCTGAGGCTT 2340
DB 2281 GCCACGCGGAGCTGCTGAGGCTTCTGAGGCTTCTGAGGCTTCTGAGGCTTCTGAGGCTT 2340
QY 2341 GCGCGGAGGCTGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
DB 2341 GCGCGGAGGCTGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
QY 2401 ATGACGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
DB 2401 ATGACGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
QY 2461 CTGCACTGGTGGATTCAGAAAGTTCAGAAAGTTCAGAAAGTTCAGAAAGTTCAGAAAG 2520
DB 2461 CTGCACTGGTGGATTCAGAAAGTTCAGAAAGTTCAGAAAGTTCAGAAAGTTCAGAAAG 2520
QY 2521 TTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
DB 2521 TTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
QY 2581 GACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
DB 2581 GACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
```

RESULT 2

ABN86735

ID ABN86735 standard; cDNA; 2640 BP.

XX AC ABN86735;

XX DT 16-SEP-2002 (first entry)

XX DE Human OC-116KDa cDNA.

XX KW Human; osteoclast; gene therapy; aberrant bone resorption; OC-116KDa;

XX KW Gene; ss.

XX OS Homo sapiens.

XX FH Key

XX FT CDS

XX FT 58..2523

XX FT /tag= a

XX FT /product= "OC-116KDa"

XX FT /note= "No stop codon given"

XX PN US6403304-B1.

XX

PD 11-JUN-2002.

XX 19-JUL-1996; 96US-00684932.

XX 06-APR-1993; 93US-00045270.

XX 23-FEB-1995; 95US-00392678.

XX 20-JUL-1995; 95US-0001292P.

XX 22-FEB-1996; 96US-000605378.

XX (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.

XX Stashenko P, Li Y, Wucherpfennig AL;

XX WPI; 2002-536031/57.

XX P-PSDB; AB881811.

XX Novel isolated osteoclast-specific or -related DNA sequence, useful for

XX producing gene products useful in the therapeutic treatment or diagnosis

XX of disorders involving aberrant bone resorption.

XX Example 8; Fig 3; 34pp; English.

XX The invention relates to novel human osteoclast-specific or -related cDNA

XX sequences. The sequence encodes human osteoclast 116KDa (OC-116KDa). The

XX sequences may have a use in gene therapy. The sequences of the invention

XX are useful in the production of gene products useful in the therapeutic

XX treatment or diagnosis of disorders involving aberrant bone resorption,

XX for generating peptides which are useful for producing antibodies for

XX identifying osteoclast-specific or -related peptides or gene products

XX Sequence 2640 BP; 445 A; 885 C; 824 G; 486 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 2640; DB 6; Length 2640;

XX Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCGTGGCGGACGCGGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60

DB 1 CGGCGTGGCGGACGCGGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60

QY 61 GGTCTCATGTTCCGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120

DB 61 GGTCTCATGTTCCGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120

QY 121 GCTACACCTGCTGAGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 180

DB 121 GCTACACCTGCTGAGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 180

QY 181 GCTCGGTGAGCGCTTCCAGAGACGCTTGTGGTGTGTGGTGTGTGGTGTGTGGTGTGTGG 240

DB 181 GCTCGGTGAGCGCTTCCAGAGACGCTTGTGGTGTGTGGTGTGTGGTGTGTGGTGTGTGG 240

QY 241 GAGAGACCTTACCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300

DB 241 GAGAGACCTTACCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300

QY 301 CCAAGGGGAGGCTGCGGACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCAC 360

DB 301 CCAAGGGGAGGCTGCGGACCCACCCACCCACCCACCCACCCACCCACCCACCCACCCAC 360

QY 361 GAGCGCTGGCGGACGAGCTGGGATGTGGGATGTGGGATGTGGGATGTGGGATGTGGGATGT 420

DB 361 GAGCGCTGGCGGACGAGCTGGGATGTGGGATGTGGGATGTGGGATGTGGGATGTGGGATGT 420

QY 421 GTGCAACGCTGAGCTCCACCGCGGCTGTACGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 480

DB 421 GTGCAACGCTGAGCTCCACCGCGGCTGTACGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 480

QY 481 GCGCGCCACACAGATGGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540

DB 481 GCGCGCCACACAGATGGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540

QY 541 CACAGGACCTGAGGCTCAACTTTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600

Db 541 CACACAGGACCTTGGAGGTCAACTTTGTGGCAGGTGCGTGGAGCCCAACAAGGCCCTCGCC 600
Qy 601 CTAGAGCGCTGCTCTGGAGGCGCTGCGCGGCTTCTCATTTGCCAGTCTCAGGAGCTG 660
Db 601 CTAGAGCGCTGCTCTGGAGGCGCTGCGCGGCTTCTCATTTGCCAGTCTCAGGAGCTG 660
Qy 661 GAGCAGCGCTGAGGACACCCCGTGAACGCGGAGCAGCAGCAGTGTGATGACCTTCTCATC 720
Db 661 GAGCAGCGCTGAGGACACCCCGTGAACGCGGAGCAGCAGCAGTGTGATGATGACCTTCTCATC 720
Qy 721 TCTACTAGGCTGAGGACAGATCGGACAGAGATCGGACAGATCACCGAGTCTTCACTGC 780
Db 721 TCTACTAGGCTGAGGACAGATCGGACAGAGATCGGACAGATCACCGAGTCTTCACTGC 780
Qy 781 CACGTCCTCCGCTTCTGACAGAGGAGGCGCGCTCGGCGCTTCCAGAGTCACTGC 840
Db 781 CACGTCCTCCGCTTCTGACAGAGGAGGCGCGCTCGGCGCTTCCAGAGTCACTGC 840
Qy 841 CAGCAGAGCAGGAGCTGACGAGGTCTCTCGGAGAGACAGAGCGGTCTCTGAGCCAGGTG 900
Db 841 CAGCAGAGCAGGAGCTGACGAGGTCTCTCGGAGAGACAGAGCGGTCTCTGAGCCAGGTG 900
Qy 901 CTAGCCCGGCTGCTGAGCTGCTGCGCAGGCGAGGTGCAAGTCAACAAGATGAAGGCC 960
Db 901 CTAGCCCGGCTGCTGAGCTGCTGCGCAGGCGAGGTGCAAGTCAACAAGATGAAGGCC 960
Qy 961 GTGTACCTGGCCCTGAAACAGTGCAGCTGAGCAGCAGCAGTGTCTCATTTCCGAG 1020
Db 961 GTGTACCTGGCCCTGAAACAGTGCAGCTGAGCAGCAGCAGTGTCTCATTTCCGAG 1020
Qy 1021 GCCTGCTGCTGTGCGAGACCTGCGCCCTGACAGGAGCCCTGCGGACAGCTCGATG 1080
Db 1021 GCCTGCTGCTGTGCGAGACTGCGCCCTGCGAGGAGCCCTGCGGACAGCTCGATG 1080
Qy 1081 GAGGAGGAGTGAAGTGCCTGCTCAACGATCCCTGCGGACATGCGCCCGCACATC 1140
Db 1081 GAGGAGGAGTGAAGTGCCTGCTCAACGATCCCTGCGGACATGCGCCCGCACATC 1140
Qy 1141 ATCCGACACCAACCGCTTCAAGCGGAGCTTCAAGGCGATCGTGATGCGTACGCGTGGC 1200
Db 1141 ATCCGACACCAACCGCTTCAAGCGGAGCTTCAAGGCGATCGTGATGCGTACGCGTGGC 1200
Qy 1201 CGCTACAGGAGGTCAACCCCGCTCCCTACACATCATCATCTTCCCTTCTGTTGCT 1260
Db 1201 CGCTACAGGAGGTCAACCCCGCTCCCTACACATCATCATCTTCCCTTCTGTTGCT 1260
Qy 1261 GTGATGTTGGGAGTGTGGGCAACGCGCTGCTCATGTTCTTCCGCTTGGCCATGCTC 1320
Db 1261 GTGATGTTGGGAGTGTGGGCAACGCGCTGCTCATGTTCTTCCGCTTGGCCATGCTC 1320
Qy 1321 CTTGGGAGAACCGACCGCTGTGAAGCGCGCAGAACGAGATCTGGCAGACTTCTTC 1380
Db 1321 CTTGGGAGAACCGACCGCTGTGAAGCGCGCAGAACGAGATCTGGCAGACTTCTTC 1380
Qy 1381 AGGCGCGCTACTGCTGCTGTATGAGCCTGTTCTCATCTACACCGGCTTCACTAC 1440
Db 1381 AGGCGCGCTACTGCTGCTGTATGAGCCTGTTCTCATCTACACCGGCTTCACTAC 1440
Qy 1441 AACGAGTGTTCAGTTCGCGCACAAGCATCTTCCCTCGGCTGAGTGTGGCGCCATG 1500
Db 1441 AACGAGTGTTCAGTTCGCGCACAAGCATCTTCCCTCGGCTGAGTGTGGCGCCATG 1500
Qy 1501 GCCAACCACTGTGGTGGAGTATGATTCCTGGGCGCAGCACCATGCTTACCTGGAT 1560
Db 1501 GCCAACCACTGTGGTGGAGTATGATTCCTGGGCGCAGCACCATGCTTACCTGGAT 1560
Qy 1561 CCCAACGTCACCGGTGTCTTCTGGGACCTTACCCCTTTGGCATCGATCTTATTTGGAGC 1620
Db 1561 CCCAACGTCACCGGTGTCTTCTGGGACCTTACCCCTTTGGCATCGATCTTATTTGGAGC 1620
Qy 1621 CTGGCTGCCAACCTTGAAGTTCCTCAACTCTTCAAGATGAAGATGTCCGTCACTC 1680

Db 1621 CTGGCTGCCAACCACTTGAAGTTCCTCAACTCTTCAAGATGAAGATGTCCGTCACTCCTG 1680
Qy 1681 GGCGTCTGTCACATGGCCCTTTGGGGTGGTCTCTCGAGTCTTCAACCAAGTGCACCTTGGC 1740
Db 1681 GGCGTCTGTCACATGGCCCTTTGGGGTGGTCTCTCGAGTCTTCAACCAAGTGCACCTTGGC 1740
Qy 1741 CAGAGGACACCGCTGCTGTCGAGACGCTGCGGAGAGTCACTTCTCTCTGGAGTCTTTC 1800
Db 1741 CAGAGGACACCGCTGCTGTCGAGACGCTGCGGAGAGTCACTTCTCTCTGGAGTCTTTC 1800
Qy 1801 GGTACTCTGCTGTTCTTAGTCACTCAAGTGGTGTGTGTGGCTGCGGAGTCTTTC 1860
Db 1801 GGTACTCTGCTGTTCTTAGTCACTCAAGTGGTGTGTGTGGCTGCGGAGTCTTTC 1860
Qy 1861 TCGCCACAGATCTCATCACTTCAATCAATGTTCTTCTTCTCCACAGCCCGCAGCAAC 1920
Db 1861 TCGCCACAGATCTCATCACTTCAATCAATGTTCTTCTTCTCCACAGCCCGCAGCAAC 1920
Qy 1921 AGGCTGCTTCTACCCCGGACAGAGTGTTCAGGCGACGCTGGTGGTCTTGGCTTGGCC 1980
Db 1921 AGGCTGCTTCTACCCCGGACAGAGTGTTCAGGCGACGCTGGTGGTCTTGGCTTGGCC 1980
Qy 1981 ATGTGCTCCTCTGCTGCTTGGCACACCTCTGCACTGCTGCAACCGCCGCGC 2040
Db 1981 ATGTGCTCCTCTGCTGCTTGGCACACCTCTGCACTGCTGCAACCGCCGCGC 2040
Qy 2041 CTGGGAGGAGCGGCTGACCGACAGAGAGAAACAGGCGGGTCTTGGACCTTGGCT 2100
Db 2041 CTGGGAGGAGCGGCTGACCGACAGAGAGAAACAGGCGGGTCTTGGACCTTGGCT 2100
Qy 2101 GACGCATCTGTGAATGGCTGAGCTCCGATCAGAGAAAGGAGGCGGCTTGGATGATAA 2160
Db 2101 GACGCATCTGTGAATGGCTGAGCTCCGATCAGAGAAAGGAGGCGGCTTGGATGATAA 2160
Qy 2161 GAGGAGCGGAGCTGCTCCCTCCGAGGTCTCATGCAACGAGGCGATCCACACCATCGAG 2220
Db 2161 GAGGAGCGGAGCTGCTCCCTCCGAGGTCTCATGCAACGAGGCGATCCACACCATCGAG 2220
Qy 2221 TTCTGCTGCTGGCTGGCTTCCAAACACCGCTTCTACCTGCGGCTGTGGGCGCTGAGCCTG 2280
Db 2221 TTCTGCTGCTGGCTGGCTTCCAAACACCGCTTCTACCTGCGGCTGTGGGCGCTGAGCCTG 2280
Qy 2281 GCCACGCGCCAGCTGTCCGAGGTCTGTGGGCGCATGTGTGATGCGCATAGGCTTGGGCGCTG 2340
Db 2281 GCCACGCGCCAGCTGTCCGAGGTCTGTGGGCGCATGTGTGATGCGCATAGGCTTGGGCGCTG 2340
Qy 2341 GCGCGGAGGTGGGCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
Db 2341 GCGCGGAGGTGGGCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
Qy 2401 ATGACCGTGGCTATCTGCTGCTGATGAGGAGTCTCAGGCTTCTCAGCGCTTCTGCGCCTGCGG 2460
Db 2401 ATGACCGTGGCTATCTGCTGCTGATGAGGAGTCTCAGGCTTCTCAGCGCTTCTGCGCCTGCGG 2460
Qy 2461 CTGCACTGGGTGGAATTCAGAACAGTTCCTACTCAGGCGAGGCTTCAAGCTGAGTCCC 2520
Db 2461 CTGCACTGGGTGGAATTCAGAACAGTTCCTACTCAGGCGAGGCTTCAAGCTGAGTCCC 2520
Qy 2521 TTGACCTTGGCTGCGACAGATGATGAGGCGCCTGAGGCTCTGCGCAGACTTCTTCTTCT 2580
Db 2521 TTGACCTTGGCTGCGACAGATGATGAGGCGCCTGAGGCTCTGCGCAGACTTCTTCTTCT 2580
Qy 2581 GACCTCTGAGGCGAGGAGGAATAAGACGCTGCGCCCTGCGCAAAAAA 2640
Db 2581 GACCTCTGAGGCGAGGAGGAATAAGACGCTGCGCCCTGCGCAAAAAA 2640

RESULT 3

ADRA4118

ID ADRA4118 standard; cDNA; 2640 BP.

XX AC ADRA4118;

XX

DT 04-NOV-2004 (first entry)
XX Human 116-kDa osteoclast proton pump subunit (OC-116 kDa) protein cDNA.
XX
KW Human; OC-116 kDa; 116-kDa osteoclast proton pump subunit; therapy;
XX bone degradation; osteoporosis; osteoarthritis; ss; gene.
XX
OS Homo sapiens.
XX
FH Key
FT CDS
FT Location/Qualifiers
FT 58..2523
FT /tag= a
FT /product= "Human 116-kDa osteoclast proton pump subunit
FT (OC-116 kDa) protein"
FT /partial
FT /note= "No stop codon"
XX
XX US6777537-B1.
XX
XX 17-AUG-2004.
XX
XX 18-JUL-2000; 2000US-00618304.
XX
XX 22-FEB-1996; 96US-00605378.
XX
XX (FOR-S-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.
XX
XX Stashenko P, Li Y;
XX WPI; 2004-591304/57.
XX P-PSDB; ADR44119.
XX
XX New isolated human 116-kDa osteoclast (OC-116 kDa) proton pump subunit
XX polypeptide, useful for preventing and/or treating diseases with bone
XX degradation, such as osteoporosis and osteoarthritis.
XX
XX Claim 1; SEQ ID NO 1; 16pp; English.
XX
XX The present invention relates to a human 116-kDa osteoclast proton pump
XX subunit (OC-116 kDa) polypeptide and its encoding polynucleotide. The
XX invention is useful for the prevention and treatment of diseases or
XX conditions associated with aberrant expression or activity of the OC-116
XX kDa protein and bone degradation, such as osteoporosis and
XX osteoarthritis. The present sequence is human OC-116 kDa protein encoding
XX cDNA.
XX
SQ Sequence 2640 BP; 445 A; 885 C; 824 G; 486 T; 0 U; 0 Other;
Query Match 100.0%; Score 2640; DB 13; Length 2640;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCGCTGCGCGGACGCCAGCAGCAGCGGAGCGCGGCGCAGCACACCGGGGACCATG 60
DB 1 CGCGCTGCGCGGACGCCAGCAGCAGCGGAGCGCGGCGCAGCACACCGGGGACCATG 60
QY 61 GGGTCCATGTTCCGAGAGGAGGAGTGCGCCTGGTCCAGCTTTTCTGCCCCACAGCGGCT 120
DB 61 GGGTCCATGTTCCGAGAGGAGGAGTGCGCCTGGTCCAGCTTTTCTGCCCCACAGCGGCT 120
QY 121 GCTTACACCTGGTGGTGGCTGGCGGAGCTGGGCTCGTGGAGTTCAGAGACCTCAAC 180
DB 121 GCTTACACCTGGTGGTGGCTGGCGGAGCTGGGCTCGTGGAGTTCAGAGACCTCAAC 180
QY 181 GCTTGGTGGAGCGCTTCCAGAGACCTTTGTTGATGTTTGGCGCTGTGAGGAGCTG 240
DB 181 GCTTGGTGGAGCGCTTCCAGAGACCTTTGTTGATGTTTGGCGCTGTGAGGAGCTG 240
QY 241 GAGAAGACCTTCACTTCTGAGAGGAGGAGTGCGCGGCTGGTGGTCTGCCCCCG 300
DB 241 GAGAAGACCTTCACTTCTGAGAGGAGGAGTGCGCGGCTGGTGGTCTGCCCCCG 300
QY 301 CCAAGGGGAGGCTGCCGGCACCCCCACCGCGGACCTGCTGGCATCCAGGAGGAGC 360

DB 301 CCAAGGGGAGGCTGCCGGCACCCCCACCGCGGACCTGCTGGCATCCAGGAGGAGC 360
QY 361 GAGCGCTTGGCCAGGAGCTGCGGATGTGCGGGCAACAGCAGGCGCTTCCGGGCCAG 420
DB 361 GAGCGCTTGGCCAGGAGCTGCGGATGTGCGGGCAACAGCAGGCGCTTCCGGGCCAG 420
QY 421 CTGACACAGCTGAGCTTCAACCGCGCTGTACGCGAGGCGCATGAACCTCAGCTGCA 480
DB 421 CTGACACAGCTGAGCTTCAACCGCGCTGTACGCGAGGCGCATGAACCTCAGCTGCA 480
QY 481 GCGCGCCACACAGATGGGCTCAGAGAGACCCCTGCTCCAGGCGCCCGGGGGCGG 540
DB 481 GCGCGCCACACAGATGGGCTCAGAGAGACCCCTGCTCCAGGCGCCCGGGGGCGG 540
QY 541 CACAGGACCTGAGGCTCAACTTTGTGGCAGGTGCGGTGGAGCCCAAGGCGCCCTGCC 600
DB 541 CACAGGACCTGAGGCTCAACTTTGTGGCAGGTGCGGTGGAGCCCAAGGCGCCCTGCC 600
QY 601 CTAGAGCGCTGCTGAGGCGCTGCGCGGCTTCTCATTTGCGAGCTTTCAGGAGCTG 660
DB 601 CTAGAGCGCTGCTGAGGCGCTGCGCGGCTTCTCATTTGCGAGCTTTCAGGAGCTG 660
QY 661 GAGCAGCGCTGAGCAGCAGCGGTGAGCGGCGAGCAGCAGCAGCTGCTTCTCATC 720
DB 661 GAGCAGCGCTGAGCAGCAGCGGTGAGCGGCGAGCAGCAGCAGCTGCTTCTCATC 720
QY 721 TCCTACTGGGCTGAGCAGATCGGACAGATCCGCAAGATCAACGAGCTTCTTCCACTGC 780
DB 721 TCCTACTGGGCTGAGCAGATCGGACAGATCCGCAAGATCAACGAGCTTCTTCCACTGC 780
QY 781 CAGCTTCCGCTTCTGAGGAGGAGCGCGCTGCGGGCCCTGCGAGCAGCTGCA 840
DB 781 CAGCTTCCGCTTCTGAGGAGGAGCGCGCTGCGGGCCCTGCGAGCAGCTGCA 840
QY 841 CAGCAGCAGGAGCTGAGGAGGCTGCGGGGAGCAGAGCGGTTCTTGAGCAGGCTG 900
DB 841 CAGCAGCAGGAGCTGAGGAGGCTGCGGGGAGCAGAGCGGTTCTTGAGCAGGCTG 900
QY 901 CTAGGCGGCTGCTGAGCTGCTGCGCCAGGAGGAGCTGAGCTTCCAGAGATGAAGGCC 960
DB 901 CTAGGCGGCTGCTGAGCTGCTGCGCCAGGAGGAGCTGAGCTTCCAGAGATGAAGGCC 960
QY 961 GTGTACTGCGCTGAAACAGTGCAGCTGAGCAGCAGCAAGTGCCTCATTTGCCGAG 1020
DB 961 GTGTACTGCGCTGAAACAGTGCAGCTGAGCAGCAGCAAGTGCCTCATTTGCCGAG 1020
QY 1021 GCTGTGCTGCTGCGAGACCTGCGCGCTGCGAGGAGCGCTGCGGGACAGCTCGATG 1080
DB 1021 GCTGTGCTGCTGCGAGACCTGCGCGCTGCGAGGAGCGCTGCGGGACAGCTCGATG 1080
QY 1081 GAGGAGGAGTGAAGTCCGCTGCTCAACCGCATCCCTGCGGGACATGCCCCCAGCCTC 1140
DB 1081 GAGGAGGAGTGAAGTCCGCTGCTCAACCGCATCCCTGCGGGACATGCCCCCAGCCTC 1140
QY 1141 ATCCGACCAACCGCTTCAACCGCTTCAACCGCTTCAACCGCTTCAACCGCTTCAAC 1200
DB 1141 ATCCGACCAACCGCTTCAACCGCTTCAACCGCTTCAACCGCTTCAACCGCTTCAAC 1200
QY 1201 CGCTACCGAGGAGTCAACCGCTTCAACCGCTTCAACCGCTTCAACCGCTTCAACCGCT 1260
DB 1201 CGCTACCGAGGAGTCAACCGCTTCAACCGCTTCAACCGCTTCAACCGCTTCAACCGCT 1260
QY 1261 GTGATGTTCCGGGATGTGGGCGCACCGGCTGCTCATGTTCTTTCGCGCTTGGCCATG 1320
DB 1261 GTGATGTTCCGGGATGTGGGCGCACCGGCTGCTCATGTTCTTTCGCGCTTGGCCATG 1320
QY 1321 CTTTGGGAGAACCGAGCGGCTGTGAAGCGCGGAGAACCGAGATCTGGCAGATTTCTTC 1380
DB 1321 CTTTGGGAGAACCGAGCGGCTGTGAAGCGCGGAGAACCGAGATCTGGCAGATTTCTTC 1380
QY 1381 AGGGGCGCTTACCTGCTTATGGGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440

| | | | |
|------|----|--|------|
| 1381 | DB | AGGGGCGCTACCTGCTCTCTGCTCTATAGGGGCGCTGTCTTCCATCTACACCGGCTTCATCTAC | 1444 |
| 1441 | QY | AACGAGTGTCTTCAGTTCGGCGCCACACAGCATCTTCCCTCTGGGCTGGAGTGTGGCGCCCATG | 1500 |
| 1441 | DB | | |
| 1441 | DB | AACGAGTGTCTTCAGTTCGGCGCCACACAGCATCTTCCCTCTGGGCTGGAGTGTGGCGCCCATG | 1500 |
| 1501 | QY | GCCAAACAGTCTGGCTGGAGTGAATGCAATCTCTGGGCCGACGACACGATGCTTACCTCTGGAT | 1560 |
| 1501 | DB | | |
| 1501 | DB | GCCAAACAGTCTGGCTGGAGTGAATGCAATCTCTGGGCCGACGACACGATGCTTACCTCTGGAT | 1560 |
| 1561 | QY | CCCAACGTCAACGGGTCTTCTCTGGGACCTTACCCCTTTGGCATGATCTATTTGGAGC | 1620 |
| 1561 | DB | | |
| 1561 | DB | CCCAACGTCAACGGGTCTTCTCTGGGACCTTACCCCTTTGGCATGATCTATTTGGAGC | 1620 |
| 1621 | QY | CTGGCTGCCAACCACTTGGAGTTCCTCAACTCTCTCAAGATGAAGATGTCGGTCATCTCTG | 1680 |
| 1621 | DB | | |
| 1621 | DB | CTGGCTGCCAACCACTTGGAGTTCCTCAACTCTCTCAAGATGAAGATGTCGGTCATCTCTG | 1680 |
| 1681 | QY | GGCGTCTGCACATGGCGTTTGGGGTGGTCTCTGGAGTCTTCAACACGTCGACTTTGGC | 1740 |
| 1681 | DB | | |
| 1681 | DB | GGCGTCTGCACATGGCGTTTGGGGTGGTCTCTGGAGTCTTCAACACGTCGACTTTGGC | 1740 |
| 1741 | QY | CAGAGGCACCGGCTGCTCTGGAGACGCTGCGGAGTCACTTCTCTCTGGGACTCTTC | 1800 |
| 1741 | DB | | |
| 1741 | DB | CAGAGGCACCGGCTGCTCTGGAGACGCTGCGGAGTCACTTCTCTCTGGGACTCTTC | 1800 |
| 1801 | QY | GTTTACCTCTGTGTTCTAGTCACTACAGTGGCTGTGTCTGGGCTGCCAGGGCGCC | 1860 |
| 1801 | DB | | |
| 1801 | DB | GTTTACCTCTGTGTTCTAGTCACTACAGTGGCTGTGTCTGGGCTGCCAGGGCGCC | 1860 |
| 1861 | QY | TCGCCCCAGCATCTCATCCACTTCATCAACATGTTCTCTTCTCCACAGCCCCAGCAAC | 1920 |
| 1861 | DB | | |
| 1861 | DB | TCGCCCCAGCATCTCATCCACTTCATCAACATGTTCTCTTCTCCACAGCCCCAGCAAC | 1920 |
| 1921 | QY | AGGCTGTCTACCCCCGGCAGGAGTGGTCCAGGCCACGCTGGTGTCTTGGGCTTGGCC | 1980 |
| 1921 | DB | | |
| 1921 | DB | AGGCTGTCTACCCCCGGCAGGAGTGGTCCAGGCCACGCTGGTGTCTTGGGCTTGGCC | 1980 |
| 1981 | QY | ATGTGGCCCATCTGTGCTTGGCACACCCTGCACCTGTGCACCGCCACACCGCCGCGC | 2040 |
| 1981 | DB | | |
| 1981 | DB | ATGTGGCCCATCTGTGCTTGGCACACCCTGCACCTGTGCACCGCCACACCGCCGCGC | 2040 |
| 2041 | QY | CTCGGAGGAGGCCGCTGACCGACAGGAGGAAACAAAGGCCGGTGTCTGGACCTTGCCT | 2100 |
| 2041 | DB | | |
| 2041 | DB | CTCGGAGGAGGCCGCTGACCGACAGGAGGAAACAAAGGCCGGTGTCTGGACCTTGCCT | 2100 |
| 2101 | QY | GACGCATCTGTGAATGCTCGAGCTCCGATGAGGAAAGCGAGGGGCTCTGGATGATGA | 2160 |
| 2101 | DB | | |
| 2101 | DB | GACGCATCTGTGAATGCTCGAGCTCCGATGAGGAAAGCGAGGGGCTCTGGATGATGA | 2160 |
| 2161 | QY | GAGGAGCCGAGCTCTGCCCTCCGAGGTCTCATGCACAGGCGCATCCACACCATCGAG | 2220 |
| 2161 | DB | | |
| 2161 | DB | GAGGAGCCGAGCTCTGCCCTCCGAGGTCTCATGCACAGGCGCATCCACACCATCGAG | 2220 |
| 2221 | QY | TTCTGCTTGGGCTGCGTCTCCAAACACCGCTCTCTACCTGCGCCTGTGGGCGCTTGAGCCTG | 2280 |
| 2221 | DB | | |
| 2221 | DB | TTCTGCTTGGGCTGCGTCTCCAAACACCGCTCTCTACCTGCGCCTGTGGGCGCTTGAGCCTG | 2280 |
| 2281 | QY | GCCCAACCGCAGTGTCCGAGGTTCTGTGGGCCATGATGATGCGCATAGGCCCTGGGCGCTG | 2340 |
| 2281 | DB | | |
| 2281 | DB | GCCCAACCGCAGTGTCCGAGGTTCTGTGGGCCATGATGATGCGCATAGGCCCTGGGCGCTG | 2340 |
| 2341 | QY | GGCGGGAGGTGGCGGTGGGCTGTGGTCCCATCTTTGGCGCTTTGGCGGTG | 2400 |
| 2341 | DB | | |
| 2341 | DB | GGCGGGAGGTGGCGGTGGGCTGTGGTCCCATCTTTGGCGCTTTGGCGGTG | 2400 |
| 2401 | QY | ATGACCGTGGCTATCTCTGCTGGTGAATGGAGGGACTCTCAGCCTTCTCTGACGCGCCTGGG | 2460 |
| 2401 | DB | | |
| 2401 | DB | ATGACCGTGGCTATCTCTGCTGGTGAATGGAGGGACTCTCAGCCTTCTCTGACGCGCCTGGG | 2460 |
| 2461 | QY | CTGCACTGGGTGGAAATTCAGAGAACAGATTCTATCTAGGCAACGGGCTACAGCTGATGCC | 2520 |
| 2461 | DB | | |
| 2461 | DB | CTGCACTGGGTGGAAATTCAGAGAACAGATTCTATCTAGGCAACGGGCTACAGCTGATGCC | 2520 |

| | | | |
|----|------|---|------|
| Qy | 2521 | TTGACCTTCGGTGCACAGATGACTAGGGCCCACTGAGAGTCTCTGCAGACCTCCTTCCT | 2580 |
| Db | 2521 | TTACCTTCGGTGCACAGATGACTAGGGCCCACTGAGAGTCTCTGCAGACCTCCTTCCT | 2580 |
| Qy | 2581 | GACCTCTGAGCGAGGAGGAATAAGACAGGGTCCGCCCTTGGCAAAAAAAAAAAAAA | 2640 |
| Db | 2581 | GACCTCTGAGCGAGGAGGAATAAGACAGGGTCCGCCCTTGGCAAAAAAAAAAAAAA | 2640 |

RESULT 4
ADY92605
ID ADY

| | |
|----|---|
| AC | ADY92605; |
| AD | |
| XX | |
| DT | 02-JUN-2005 (first entry) |
| XX | |
| DE | Human 116 kD osteoclast proton pump OC-116 subunit cDNA. |
| XX | |
| KW | cloning; osteopathic; antiarthritic; gene therapy; |
| KW | osteoclast proton pump; bone disease; osteoporosis; osteoarthritis; gene; |
| XX | ss. |
| KW | |
| XX | |
| OS | Homo sapiens. |
| XX | |
| FH | Key Location/Qualifiers |
| FT | 58..2547 |
| FT | /*tag= a |
| FT | /product= "116 kD osteoclast proton pump subunit" |
| XX | |
| FN | US2005064448-A1. |
| XX | |
| PD | 24-MAR-2005. |
| XX | |
| PF | 20-FEB-2004; 2004US-00783519. |
| XX | |
| PK | 22-FEB-1996; 96US-00605378. |
| PK | 18-JUL-2000; 2000US-00618304. |
| XX | |
| PA | (STAS/) STASHENKO P. |
| PA | (LIYY/) LI Y. |
| XX | |
| PI | Stashenko P, Li Y; |
| XX | |
| DR | WPI; 2005-241271/25. |
| XX | P-PSDB; ADY92606. |
| XX | |
| PT | New isolated genes and encoded human 116-kDa osteoclast proton pump |
| PT | subunit polypeptides useful for treating bone mass disorders associated |
| PT | with aberrant rate of bone degradation (e.g. osteoporosis or |
| XX | osteoarthritis). |
| XX | |
| PS | Claim 3; SEQ ID NO 1; 17pp; English. |
| XX | |
| CC | The invention relates to an isolated gene encoding a polypeptide which is |
| CC | a human 116-kD osteoclast proton pump subunit. The composition and |
| CC | methods are useful for treating bone mass disorders characterized by |
| CC | aberrant rate of bone degradation (e.g. osteoporosis or osteoarthritis). |
| CC | The protein may also be used as cell surface markers for osteoclasts. |
| CC | This sequence corresponds to the cDNA encoding the novel proton pump |
| CC | subunit. The cDNA was isolated by differential screening of a human |
| CC | osteoclastoma cDNA library. |
| XX | |
| SQ | Sequence 2640 BP; 445 A; 885 C; 824 G; 486 T; 0 U; 0 Other; |

| Query Match | 100.0%; | Score 2640; | DB 14; | Length 2640; |
|-----------------------|-----------------|--|-----------|--------------|
| Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| Matches 2640; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 1 | CGGGCTGCGGGGACGGGCGAGCAGCAGCGAGGCGGGCGGCGAGCACACCCGGGGACCATG | 60 | |
| Db | 1 | CGGGCTGCGGGGACGGGCGAGCAGCAGCGAGGCGGGCGGCGAGCACACCCGGGGACCATG | 60 | |

| | | | |
|----|------|---|------|
| Qy | 241 | GAGAGACCTTCACTTCTCTCAGGAGGAGGTGCGCGGGCTGGGCTGGTCTCTGCCCCCG | 300 |
| Db | 241 | GAGAGACCTTCACTTCTCTCAGGAGGAGGTGCGCGGGCTGGGCTGGTCTCTGCCCCCG | 300 |
| Qy | 301 | CCAAAGGGGAGGCTGCCGGCACCCCGGACCTGCTGCGCATCTCCAGGAGGAGACG | 360 |
| Db | 301 | CCAAAGGGGAGGCTGCCGGCACCCCGGACCTGCTGCGCATCTCCAGGAGGAGACG | 360 |
| Qy | 361 | GAGCGCTTGGCCACAGAGCTGCGGGATGTGCGGGCAACACAGCAGGCCCTCTCGGGCCCCAG | 420 |
| Db | 361 | GAGCGCTTGGCCACAGGAGCTGCGGGATGTGCGGGCAACACAGCAGGCCCTCTCGGGCCCCAG | 420 |
| Qy | 421 | CTGCACCACTGCAGCTCCACGCGCGGTCTACGCCAGGGCCATGAACCTTCAGCTGGCA | 480 |
| Db | 421 | CTGCACCACTGCAGCTCCACGCGCGGTCTACGCCAGGGCCATGAACCTTCAGCTGGCA | 480 |
| Qy | 481 | GCCGCCCAACAGATGGGGCTCAGAGAGCAGCCCTGCTTCCAGGGCCCCCGGGGGCCG | 540 |
| Db | 481 | GCCGCCCAACAGATGGGGCTCAGAGAGCAGCCCTGCTTCCAGGGCCCCCGGGGGCCG | 540 |
| Qy | 541 | CACAGGACCTGAGGGTCAACTTTGTGGCAGTGTCGTGAGGCCCAACAAGGCCCTTGCC | 600 |
| Db | 541 | CACAGGACCTGAGGGTCAACTTTGTGGCAGTGTCGTGAGGCCCAACAAGGCCCTTGCC | 600 |
| Qy | 601 | CTAGAGCGCTGCTCTGAGGGCCCTGCGGGCTTCTCATGTCAGCTTCAGGAGCTG | 660 |
| Db | 601 | CTAGAGCGCTGCTCTGAGGGCCCTGCGGGCTTCTCATGTCAGCTTCAGGAGCTG | 660 |
| Qy | 661 | GAGCAGCGCTGAGCAGCCCCGTGACGGGGCAGCAGCCAGTGGATGAACCTTCTCATC | 720 |
| Db | 661 | GAGCAGCGCTGAGCAGCCCCGTGACGGGGCAGCAGCCAGTGGATGAACCTTCTCATC | 720 |
| Qy | 721 | TCCTACTGGGGTGAAGATCGGACAGAGATCCGCAAGATCAGGACTGTTCTCACTGC | 780 |
| Db | 721 | TCCTACTGGGGTGAAGATCGGACAGAGATCCGCAAGATCAGGACTGTTCTCACTGC | 780 |
| Qy | 781 | CACGTCTTCCGTTTCTGACGAGGAGGAGCCCGCTCGGGGCCCTGACGAGCTGCA | 840 |
| Db | 781 | CACGTCTTCCGTTTCTGACGAGGAGGAGCCCGCTCGGGGCCCTGACGAGCTGCA | 840 |
| Qy | 841 | CAGCAGAGCCAGGAGCTGCAGGAGGTCTCTCGGGAGACAGAGCGGTTCTTGAGCCAGGTG | 900 |
| Db | 841 | CAGCAGAGCCAGGAGCTGCAGGAGGTCTCTCGGGAGACAGAGCGGTTCTTGAGCCAGGTG | 900 |
| Qy | 901 | CTAGGCGGGTGTGACAGCTGCTGCCCGCAGGGCAGGTGCAGGTCCACAGATGAAGGCC | 960 |
| Db | 901 | CTAGGCGGGTGTGACAGCTGCTGCCCGCAGGGCAGGTGCAGGTCCACAGATGAAGGCC | 960 |
| Qy | 961 | GTGTACTTGGCCCTGAACCACTGAGCAGCAGCAGCAAGTGCCTCATTTGCCGAG | 1020 |
| Db | 961 | GTGTACTTGGCCCTGAACCACTGAGCAGCAGCAGCAAGTGCCTCATTTGCCGAG | 1020 |
| Qy | 1021 | GCCTGTGCTCTGTGCGAGACCTGCCCGCTGACGAGGCCCTTGCGGGAAGCTCGATG | 1080 |
| Db | 1021 | GCCTGTGCTCTGTGCGAGACCTGCCCGCTGACGAGGCCCTTGCGGGAAGCTCGATG | 1080 |
| Qy | 1081 | GAGGAGGAGTGAGTGCCTGGGTCTACCGCATCCCTCGCGGACATGCCCCCAACATC | 1140 |
| Db | 1081 | GAGGAGGAGTGAGTGCCTGGGTCTACCGCATCCCTCGCGGACATGCCCCCAACATC | 1140 |
| Qy | 1141 | ATCCGACCAACCGCTTTCAGGGCCAGTTCACAGGGCATCGTGGATGCTACGCGCTGGGC | 1200 |
| Db | 1141 | ATCCGACCAACCGCTTTCAGGGCCAGTTCACAGGGCATCGTGGATGCTACGCGCTGGGC | 1200 |
| Qy | 1201 | CGCTACCGAGAGTCAACCCCGCTTCCCTACACCATATCACTTCCCTTCTGTTTGTCT | 1260 |
| Db | 1201 | CGCTACCGAGAGTCAACCCCGCTTCCCTACACCATATCACTTCCCTTCTGTTTGTCT | 1260 |
| Qy | 1261 | GTGATGTTCCGGGATGTGGGCCACGGGCTGCTCATGTTCTCTTCGCCCTTGCCATGCTC | 1320 |
| Db | 1261 | GTGATGTTCCGGGATGTGGGCCACGGGCTGCTCATGTTCTCTTCGCCCTTGCCATGCTC | 1320 |

| | | | |
|----|------|--|------|
| Qy | 1321 | CTTCCGAGAAACCGACCGCTGTGAAGCCGCGCAGAAACGAGATCTGGCAGACTTTCTTC | 1380 |
| Db | 1321 | CTTTCGAGAAACCGACCGCTGTGAAGCCGCGCAGAAACGAGATCTGGCAGACTTTCTTC | 1380 |
| Qy | 1381 | AGGGCCGCTACCTGCTTCCTGCTTATGGGCTGCTTCTCATCTPACACCGGCTTCATCTAC | 1440 |
| Db | 1381 | AGGGCCGCTACCTGCTTCCTGCTTATGGGCTGCTTCTCATCTPACACCGGCTTCATCTAC | 1440 |
| Qy | 1441 | AACGAGTGTTCAGTCCGCGCCACAGCATCTTCCCCTCGGGCTGGAGTGTGGCCGCCATG | 1500 |
| Db | 1441 | AACGAGTGTTCAGTCCGCGCCACAGCATCTTCCCCTCGGGCTGGAGTGTGGCCGCCATG | 1500 |
| Qy | 1501 | GCCAAACAGTCTGGCTGGAGTGATGATTCCTGGCCACGACACAGATGCTTACCTGGAT | 1560 |
| Db | 1501 | GCCAAACAGTCTGGCTGGAGTGATGATTCCTGGCCACGACACAGATGCTTACCTGGAT | 1560 |
| Qy | 1561 | CCCAAACCTACCGGTGTCTTCTGGGACCTACCCCTTTGGGATCGATCGATTTTGGAGC | 1620 |
| Db | 1561 | CCCAAACCTACCGGTGTCTTCTGGGACCTACCCCTTTGGGATCGATCGATTTTGGAGC | 1620 |
| Qy | 1621 | CTGGCTGCCAAACCACTTGAGCTTCCTCAACTCTTCAAGATGAAGATGTCGCTCACTTC | 1680 |
| Db | 1621 | CTGGCTGCCAAACCACTTGAGCTTCCTCAACTCTTCAAGATGAAGATGTCGCTCACTTC | 1680 |
| Qy | 1681 | GGCGTGGTGACATGGCGTTTGGGGTGGTCTCTCGAGTCTTCAACACGCTGCACTTGGC | 1740 |
| Db | 1681 | GGCGTGGTGACATGGCGTTTGGGGTGGTCTCTCGAGTCTTCAACACGCTGCACTTGGC | 1740 |
| Qy | 1741 | CAGAGGCACCGCTGCTGGAGACGCTGCGGAGCTCACCTTCTGCTGGAGCTCTTC | 1800 |
| Db | 1741 | CAGAGGCACCGCTGCTGGAGACGCTGCGGAGCTCACCTTCTGCTGGAGCTCTTC | 1800 |
| Qy | 1801 | GGTTACCTCGTGTTCCTAGTCACTCAAGTGGCTGTGTCTGGGCTGCCAGGGCCGCC | 1860 |
| Db | 1801 | GGTTACCTCGTGTTCCTAGTCACTCAAGTGGCTGTGTCTGGGCTGCCAGGGCCGCC | 1860 |
| Qy | 1861 | TCGCCAGCATCTCATCACTTCATCAACATGTTTCTTCTCCACAGCCGCCAGCAAC | 1920 |
| Db | 1861 | TCGCCAGCATCTCATCACTTCATCAACATGTTTCTTCTCCACAGCCGCCAGCAAC | 1920 |
| Qy | 1921 | AGGCTGCTTACCCCGGACGAGGTGGTCCAGGCCACGCTGGTGGTCTCGGCTTGGCC | 1980 |
| Db | 1921 | AGGCTGCTTACCCCGGACGAGGTGGTCCAGGCCACGCTGGTGGTCTCGGCTTGGCC | 1980 |
| Qy | 1981 | ATGGTGCCCATCTCTGCTTGGCACACCCCTGCACCTGTGTGACCGCCACCGCCGCCG | 2040 |
| Db | 1981 | ATGGTGCCCATCTCTGCTTGGCACACCCCTGCACCTGTGTGACCGCCACCGCCGCCG | 2040 |
| Qy | 2041 | CTCGGAGGAGGCCCGCTGACCGACAGAGGAGAAAACAGGCCGGGTGTGTGACCTTGCCT | 2100 |
| Db | 2041 | CTCGGAGGAGGCCCGCTGACCGACAGAGGAGAAAACAGGCCGGGTGTGTGACCTTGCCT | 2100 |
| Qy | 2101 | GAGCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAAGGAGGGGCTTGATGATGAA | 2160 |
| Db | 2101 | GAGCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAAGGAGGGGCTTGATGATGAA | 2160 |
| Qy | 2161 | GAGGAGGCCGAGCTCGTCCCTCCGAGGTGCTCATGCAACAGGCCCATCCACACCATCGAG | 2220 |
| Db | 2161 | GAGGAGGCCGAGCTCGTCCCTCCGAGGTGCTCATGCAACAGGCCCATCCACACCATCGAG | 2220 |
| Qy | 2221 | TTCTGCTGGGCTGCGTCTCCAAACCGGCTCTACTCTGCGCTGTGGGCCCTGAGCCTG | 2280 |
| Db | 2221 | TTCTGCTGGGCTGCGTCTCCAAACCGGCTCTACTCTGCGCTGTGGGCCCTGAGCCTG | 2280 |
| Qy | 2281 | GCCACGCCAGCTGTCCGAGTTCGTGGGCCATGGTGATGGGCATAGGCCCTGGGCCCTG | 2340 |
| Db | 2281 | GCCACGCCAGCTGTCCGAGTTCGTGGGCCATGGTGATGGGCATAGGCCCTGGGCCCTG | 2340 |
| Qy | 2341 | GGCCGGAGGTGGCGGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG | 2400 |
| Db | 2341 | GGCCGGAGGTGGCGGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG | 2400 |
| Qy | 2401 | ATGACCGTGCTATCTCTGCTGGTGATGAGGAGCACTCTCAGCCCTTCTGACCGCCCTGGG | 2460 |

| | | | | | | | | | | | |
|----|------|---------|-------------|-----------|---------|------------|------------|---------------|----------------------|------|-----|
| QY | 421 | CTGCA | CCAGCTG | CAGCTCA | CGCCG | CCGTGCTA | CGCCAGG | GCCATGA | AACCTCAGCTG | GCA | 480 |
| DB | 421 | CTGCA | CCAGCTG | CAGCTCA | CGCCG | CCCGCTGT | AGCCAGG | GCCATGA | AACCTCAGCTG | GCA | 480 |
| QY | 481 | GCCGCC | CCACACAGAT | TGGGGCTC | CAGAGAG | AGCCCTCGCT | CCAGGCC | CCCGGGGGGGCGG | CGG | 540 | |
| DB | 481 | GCCGCC | CCACACAGAT | TGGGGCTC | CAGAGAG | AGCCCTCGCT | CCAGGCC | CCCGGGGGGGCGG | CGG | 540 | |
| QY | 541 | CACCAGG | ACCTGAGGGT | CAACTTT | TGTG | CAGGTG | CCGTGAGAG | CCGCCCA | CAAGGCCCCCTTGCC | 600 | |
| DB | 541 | CACCAGG | ACCTGAGGGT | CAACTTT | TGTG | CAGGTG | CCGTGAGAG | CCGCCCA | CAAGGCCCCCTTGCC | 600 | |
| QY | 601 | CTAGAG | CGCCTGCTCT | GAGAGG | CCCTG | CCCGCTTCCT | CAATTG | CCAGCTTCAGG | AGCTG | 660 | |
| DB | 601 | CTAGAG | CGCCTGCTCT | GAGAGG | CCCTG | CCCGCTTCCT | CAATTG | CCAGCTTCAGG | AGCTG | 660 | |
| QY | 661 | GAGCAG | CCCGCTGG | AGCAC | CCCGTGA | CGGGCAG | CCCGCCAGT | CGGTGAT | CACTTCCTCATC | 720 | |
| DB | 661 | GAGCAG | CCCGCTGG | AGCAC | CCCGTGA | CGGGCAG | CCCGCCAGT | CGGTGAT | CACTTCCTCATC | 720 | |
| QY | 721 | TCCTACT | GGGGTG | AGCAGAT | CGGAC | GAGAAGAT | CCCGAAGAT | CA | CGGACTGCTTCCACTGC | 780 | |
| DB | 721 | TCCTACT | GGGGTG | AGCAGAT | CGGAC | GAGAAGAT | CCCGAAGAT | CA | CGGACTGCTTCCACTGC | 780 | |
| QY | 781 | CACGTCT | TTCCGTTCTG | CACG | AGGAGG | CCCGCTCG | GGGCGCTTCG | CAGCAGCT | GCA | 840 | |
| DB | 781 | CACGTCT | TTCCGTTCTG | CACG | AGGAGG | CCCGCTCG | GGGCGCTTCG | CAGCAGCT | GCA | 840 | |
| QY | 841 | CAGCAG | ACCCAGGACTG | CAGAGG | TCTCG | GGGAGACAG | AGCCGTTCT | CTGAGCCAG | GTG | 900 | |
| DB | 841 | CAGCAG | ACCCAGGACTG | CAGAGG | TCTCG | GGGAGACAG | AGCCGTTCT | CTGAGCCAG | GTG | 900 | |
| QY | 901 | CTAGG | CCGGGTGCTG | CAGCTCT | CGCCAG | GGCAGGTG | CAGGTCC | CAAGATGA | AGGCC | 960 | |
| DB | 901 | CTAGG | CCGGGTGCTG | CAGCTCT | CGCCAG | GGCAGGTG | CAGGTCC | CAAGATGA | AGGCC | 960 | |
| QY | 961 | GTGTACT | CTGGCCCTG | AAACAGT | GAGCGT | GAGCA | CCAGCA | AAAGTGC | CAATTGCCGAG | 1020 | |
| DB | 961 | GTGTACT | CTGGCCCTG | AAACAGT | GAGCGT | GAGCA | CCAGCA | AAAGTGC | CAATTGCCGAG | 1020 | |
| QY | 1021 | GCCTGT | GCTCTGTG | CGGAGCT | CGCCGCT | CTG | CAGAGGCC | CTCGG | GACAGCTCGATG | 1080 | |
| DB | 1021 | GCCTGT | GCTCTGTG | CGGAGCT | CGCCGCT | CTG | CAGAGGCC | CTCGG | GACAGCTCGATG | 1080 | |
| QY | 1081 | GAGGAG | GAGTGTG | CTGGCT | CACG | CATCCCTG | CGGGACAT | ATGCCCC | CACACTC | 1140 | |
| DB | 1081 | GAGGAG | GAGTGTG | CTGGCT | CACG | CATCCCTG | CGGGACAT | ATGCCCC | CACACTC | 1140 | |
| QY | 1141 | ATCCGCA | CCAA | CCGCTTCA | CGGCCAG | CTTCCAGGG | CCATCG | TGGATCG | CTACGGCGTGGGC | 1200 | |
| DB | 1141 | ATCCGCA | CCAA | CCGCTTCA | CGGCCAG | CTTCCAGGG | CCATCG | TGGATCG | CTACGGCGTGGGC | 1200 | |
| QY | 1201 | CGCTAC | CCAGGAGTCA | ACCCGCT | CTCCTA | CA | CAATCAT | CACTTCC | CCCTCTCTGTTGCT | 1260 | |
| DB | 1201 | CGCTAC | CCAGGAGTCA | ACCCGCT | CTCCTA | CA | CAATCAT | CACTTCC | CCCTCTCTGTTGCT | 1260 | |
| QY | 1261 | GTGAT | TTTCGGGGAT | GTGGG | CCAGG | CTGTCT | ATGTTCTCT | TC | TCGCCCCCTGGCCATGGTC | 1320 | |
| DB | 1261 | GTGAT | TTTCGGGGAT | GTGGG | CCAGG | CTGTCT | ATGTTCTCT | TC | TCGCCCCCTGGCCATGGTC | 1320 | |
| QY | 1321 | CTTCG | CGGAGAACCG | ACCGGCTGT | GAAAG | CCGCG | CAGAACG | AGATCT | TCGGCAGACTTCTTC | 1380 | |
| DB | 1321 | CTTCG | CGGAGAACCG | ACCGGCTGT | GAAAG | CCGCG | CAGAACG | AGATCT | TCGGCAGACTTCTTC | 1380 | |
| QY | 1381 | AGGGG | CCGCTACCTG | CTCTGCT | TATGG | CCCTGTTT | CCATCT | ATAC | CCGGTTCATCTAC | 1440 | |
| DB | 1381 | AGGGG | CCGCTACCTG | CTCTGCT | TATGG | CCCTGTTT | CCATCT | ATAC | CCGGTTCATCTAC | 1440 | |
| QY | 1441 | AACGAG | TGCTTTCAGT | TCGGCC | CAC | CAGCAT | CTTCC | CCCTCG | GGCTGGAGTGTGGCGCCATG | 1500 | |
| DB | 1441 | AACGAG | TGCTTTCAGT | TCGGCC | CAC | CAGCAT | CTTCC | CCCTCG | GGCTGGAGTGTGGCGCCATG | 1500 | |

| | | | | |
|----|------|-------|--|------|
| QY | 1501 | GCCAA | CAGTCTGCTCGAGTGAATGCAATCTCTGGCCAGACACGATGCTTACCCCTGGAT | 1560 |
| DB | 1501 | GCCAA | CAGTCTGGCTCGAGTGAATGCAATCTCTGGCCAGACACGATGCTTACCCCTGGAT | 1560 |
| QY | 1561 | CCCA | AGTGCACCGGTGCTTCCTGGGACCCCTACCCCTTTGGCATCGATCCTATTTGGAGC | 1620 |
| DB | 1561 | CCCA | AGTGCACCGGTGCTTCCTGGGACCCCTACCCCTTTGGCATCGATCCTATTTGGAGC | 1620 |
| QY | 1621 | CTGG | CTGCCAACCACTTGAGCTTCTCAA | 1680 |
| DB | 1621 | CTGG | CTGCCAACCACTTGAGCTTCTCAA | 1680 |
| QY | 1681 | GGCT | GTGCAATGCGCTTGGGGTGGTCTCAAGATGAAGATGTCCGTCACTCTTG | 1740 |
| DB | 1681 | GGCT | GTGCAATGCGCTTGGGGTGGTCTCAAGATGAAGATGTCCGTCACTCTTG | 1740 |
| QY | 1741 | CAGAG | GCACCGGTGCTGCGAGACGCTGCGGAGCTCACCTTCTGTGGGACTCTTC | 1800 |
| DB | 1741 | CAGAG | GCACCGGTGCTGCGAGACGCTGCGGAGCTCACCTTCTGTGGGACTCTTC | 1800 |
| QY | 1801 | GTTT | ACCTCGTGTTCCTAGTCACTCA | 1860 |
| DB | 1801 | GTTT | ACCTCGTGTTCCTAGTCACTCA | 1860 |
| QY | 1861 | TCG | CCAGCATCTCATCACTTCA | 1920 |
| DB | 1861 | TCG | CCAGCATCTCATCACTTCA | 1920 |
| QY | 1921 | AGG | CTGCTTACCCCGGACGAGGTGGTCCAGGCCACCGTGTGTCTTGGCCTTGGCC | 1980 |
| DB | 1921 | AGG | CTGCTTACCCCGGACGAGGTGGTCCAGGCCACCGTGTGTCTTGGCCTTGGCC | 1980 |
| QY | 1981 | ATG | GTGCCCATCTGTGCTTGGCACACCCCTGCACCTGTGACCGCACCGCCGCCCG | 2040 |
| DB | 1981 | ATG | GTGCCCATCTGTGCTTGGCACACCCCTGCACCTGTGACCGCACCGCCGCCCG | 2040 |
| QY | 2041 | CTG | CGAGGAGGCCCTGTGACCGACAGGAGGAAAA | 2100 |
| DB | 2041 | CTG | CGAGGAGGCCCTGTGACCGACAGGAGGAAAA | 2100 |
| QY | 2101 | GAC | CATCTGTAATGCTGGAGCTCCGATGAGGAAAAAGCAGGGGGCTTGNATGAA | 2160 |
| DB | 2101 | GAC | CATCTGTAATGCTGGAGCTCCGATGAGGAAAAAGCAGGGGGCTTGNATGAA | 2160 |
| QY | 2161 | GAG | GAGCCGAGCTCGTCCCTCCGAGGTGCTCATGCA | 2220 |
| DB | 2161 | GAG | GAGCCGAGCTCGTCCCTCCGAGGTGCTCATGCA | 2220 |
| QY | 2221 | TTCT | GCTGGGCTGCTCCTCAACACCGCTCTTACCTGCGCTGTGGGCCCTTGAGCCTG | 2280 |
| DB | 2221 | TTCT | GCTGGGCTGCTCCTCAACACCGCTCTTACCTGCGCTGTGGGCCCTTGAGCCTG | 2280 |
| QY | 2281 | GCC | ACGCCAGCTGTCCGAGGTCTGTGGGCCATGGTGA | 2340 |
| DB | 2281 | GCC | ACGCCAGCTGTCCGAGGTCTGTGGGCCATGGTGA | 2340 |
| QY | 2341 | GGC | CGGAGGTGGCGTGGCGGTGTGTGCTCCCATCTTTGCGCTTTGCGCGTG | 2400 |
| DB | 2341 | GGC | CGGAGGTGGCGTGGCGGTGTGTGCTCCCATCTTTGCGCGCTTTGCGCGTG | 2400 |
| QY | 2401 | ATGA | CCGTGGCTATCTGTGGTGA | 2460 |
| DB | 2401 | ATGA | CCGTGGCTATCTGTGGTGA | 2460 |
| QY | 2461 | CTGA | CTGGGTGAAATTCAGAA | 2520 |
| DB | 2461 | CTGA | CTGGGTGAAATTCAGAA | 2520 |
| QY | 2521 | TTCA | CTTCTGCA | 2580 |
| DB | 2521 | TTCA | CTTCTGCA | 2580 |
| QY | 2581 | GAC | CTTGAGGACGAGGAAATAAGACGGTCCGCGCTGGCAAAAAAAAAAAAAA | 2640 |

2581 GACCTCTGAGCGCAGGAGGAATAAGACGCTCCGCCCTCGCAAAAAAAAAAAAAAAAAAAAA 2640

Db

RESULT 7

ABK84337

ID ABK84337 standard; cDNA; 2655 BP.

XX AC ABK84337;

XX

DT 14-AUG-2002 (first entry)

XX

XX Human cDNA differentially expressed in granulocytic cells #908.

XX

XX Human; ss; granulocytic cell; DNA chip; bacterial infection;

XX KW viral infection; parasitic infection; protozoal infection;

XX KW fungal infection; sterile inflammatory disease; psoriasis;

XX KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;

XX KW cardiac reperfusion injury; renal reperfusion injury; ARDS;

XX KW adult respiratory distress syndrome; inflammatory bowel disease;

XX KW Crohn's disease; ulcerative colitis; periodontal disease;

XX KW granulocyte activation; chronic inflammation; allergy.

XX

OS Homo sapiens.

XX

XX MO200228999-A2.

XX

XX 11-APR-2002.

PD

XX

XX 03-OCT-2001; 2001MO-US030821.

PF

XX

XX 03-OCT-2000; 2000NS-0237189P.

PR

XX

XX (GENE-) GENE LOGIC INC.

PA

XX

XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;

PI

XX

XX WPI; 2002-435328/46.

DR

XX

XX

PT Detecting granulocyte activation by detecting differential expression of genes associated with granulocyte activation, which serves as diagnostic markers that is useful for monitoring disease states and drug toxicity.

PT

XX

XX Claim 1; SEQ ID NO 908; 114pp; English.

PS

XX

XX The invention relates to detecting (M1) granulocyte (GC) activation (GCA), by detecting the level of expression of gene(s) (Gs) identified by DNA chip analysis as given in the specification, and comparing the expression level to an expression level in an unactivated GC, where differential expression of Gs is indicative of GCA. Also included are modulating (M2) GA by contacting GC with an agent that alters the expression of at least one gene in Gs; (2) screening (M3) for an agent capable of modulating GCA or an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease using the gene expression profile; (3) detecting (M4) an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by detecting the level of expression in a sample of the tissue of gene(s) from Gs, where the level of expression of the gene is indicative of inflammation; (4) treating (M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease; also bacterial

WPI; 2002-426119/45.

Diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumor in a patient, involves detecting the level of expression of two or more genes in a liver tissue sample.

Claim 1; SEQ ID NO 3364; 298pp; English.

The invention relates to a novel method for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumour in a patient, and differentiating metastatic liver cancer from hepatocellular carcinoma in a patient, involving detecting the level of expression of two or more genes represented in ABN97455 in a tissue sample. The method of the invention has hepatotropic, and cytostatic activity. The method is useful for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma and metastatic liver carcinoma in a patient. The method is useful for identifying expression profiles which serve as useful diagnostic markers as well as markers that can be used to monitor disease states, disease progression, drug toxicity, drug efficacy and drug metabolism. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct_sequences/

Sequence 2655 BP; 460 A; 885 C; 824 G; 486 T; 0 U; 0 Other;

Query Match 100.0%; Score 2640; DB 6; Length 2655;

Best Local Similarity

Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGCGTGCGGACGGGCAGCCAGCAGCGGAGCGCGCGGCGGACCCGGGACCATG 60

b 1 CGCGTGGCGACGGGCAGCCAGCAGCGGAGCGCGCGCAGCACACCGGGACCATG 60

61 GGCTCCATGTTCCGGAGCGAGAGGTGGCCCTGGTCCAGCTCTTTCTGCCACAGCGGCT 120

b 61 GGTCCATGTTCCGGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCACAGCGGCT 120

121 GCCTACACCTGCGTGAGTCGGCTGGCGGAGCTGGGCCCTCGTGGAGTTCAGAGACCTCAAC 180

121 GCCTACACCTGCGTGAGTCGGCTGGCGAGCTGGCCCTCGTGAGTTCAGAGACCTCAAC 180

181 GCCTCGGTGAGCGCCTTCCAGAGACGCTTGTGGTTGATGTTTGGCGCTGTGAGGAGCTG 240

181 GCCTCGGTGAGCGCCTTCCAGAGACGCTTTGTGGTTGATGTTTGGCGCTGTGAGGAGCTG 240

241 GAGAAGACCTTCACCTTCCTGCAGGAGGAGGTGCGCGGGCTGGGCTGGTCTGCCCCCG 300

241 GAGAAGACCTTCACCTTCCTGCAGGAGGAGGTGCGCGGGCTGGGCTGGTCTGCCCCCG 300

301 CCAAGGGAGGCTGCCGGCACCCCGGACCTGCTGCCGATCCAGGAGGACG 360

301 CCAAAGGGAGGCTGCCGGCACCCCA CCCCCTGCTGGCATCCAGGAGGACG 360

361 GAGCGCTGGCCACGAGCTGCGGGATGTGCGGGCAACAGCAGGCCCTGCGGGCCAG 420

361 GAGCGCCTGGCCCAAGGAGCTGGCGGATGTGGGGGCAACGAGCAGGCGCTGGGGCCAG 420

421 CTGCAACGAGTGCAGCTCCACGCCGCCGTGCTACGCCAGGGCCATGAACCTCAGCTGGCA 480

421 CTGCA CCAGCTGCAGCTCCACGGCCGCGCTCATCGCCAGGGCCATCAACCTACGCTGCCA 480

481 GCGGCCACACAGATGGGGCTCAGAGAGGAGCGCCCTGCTCCAGGCCCCCGGGGGCGG 540

[illegible]

541 CACAGGACCTGAGGGTCAATTCTGTGGCAGGTGCGCTCGAGCCCAAGGCGCCCTGCC 500

[illegible]

607 CTTGGGCGCTTCCTTCCGCCGTATCCTCCACGCAGCACTCC
650

[illegible]

| | | | |
|----|------|--|------|
| Qy | 1741 | CAGAGGCACCGGCTGCTCTGGAGACGCTGCCGAGCTCACTTCTCTGCTGGACTCTTC | 1800 |
| Db | 1741 | CAGAGGCACCGGCTGCTCTGGAGACGCTGCCGAGCTCACTTCTCTGCTGGACTCTTC | 1800 |
| Qy | 1801 | GGTTACCTCTGTTTCTTAGTCACTACAAGTGGCTGTGTCTGGGCTGCCAGGCGGCC | 1860 |
| Db | 1801 | GGTTACCTCTGTTTCTTAGTCACTACAAGTGGCTGTGTCTGGGCTGCCAGGCGGCC | 1860 |
| Qy | 1861 | TCGCCCAGCATCTCATCCACTTCATCAACATGTTCTCTTCTCCACAGCCCCAGCAAC | 1920 |
| Db | 1861 | TCGCCCAGCATCTCATCCACTTCATCAACATGTTCTCTTCTCCACAGCCCCAGCAAC | 1920 |
| Qy | 1921 | AGGCTGCTTACCCCGCAGGAGGTGGTCCAGGCCACGCTGGTGGTCTTGGCTTGGCC | 1980 |
| Db | 1921 | AGGCTGCTTACCCCGCAGGAGGTGGTCCAGGCCACGCTGGTGGTCTTGGCTTGGCC | 1980 |
| Qy | 1981 | ATGCTGCCATCTCTGCTGTCTGGCACAACCTCTGCACCTCTGTGCAACGCCACCGCCGCGC | 2040 |
| Db | 1981 | ATGCTGCCATCTCTGCTGTCTGGCACAACCTCTGCACCTCTGTGCAACGCCACCGCCGCGC | 2040 |
| Qy | 2041 | CTCGGAGGAGGCCGCTGACCGACAGGAGGAAACAAGGCCGGTGTCTGGACCTGCCT | 2100 |
| Db | 2041 | CTCGGAGGAGGCCGCTGACCGACAGGAGGAAACAAGGCCGGTGTCTGGACCTGCCT | 2100 |
| Qy | 2101 | GAGCATCTGTGAATGSGCTGGAGCTCCGATCAGGAAAAGCAGGGGCCCTGGATGATGAA | 2160 |
| Db | 2101 | GAGCATCTGTGAATGSGCTGGAGCTCCGATCAGGAAAAGCAGGGGCCCTGGATGATGAA | 2160 |
| Qy | 2161 | GAGGAGCCGAGCTGCTCCCTCCGAGGTGCTATGCAACAGGCCATTCACACCATGAG | 2220 |
| Db | 2161 | GAGGAGCCGAGCTGCTCCCTCCGAGGTGCTATGCAACAGGCCATTCACACCATGAG | 2220 |
| Qy | 2221 | TTCTGCTGGGCTGSGCTCTCAAACACCGCTCTTACTCTGCGCTGTGGGCCCTGAGGCTG | 2280 |
| Db | 2221 | TTCTGCTGGGCTGSGCTCTCAAACACCGCTCTTACTCTGCGCTGTGGGCCCTGAGGCTG | 2280 |
| Qy | 2281 | GCCACGCCCAGCTGTCCGAGGTTCTGTGGGCCATGATGAGCGCATAGGCTTGGGCTG | 2340 |
| Db | 2281 | GCCACGCCCAGCTGTCCGAGGTTCTGTGGGCCATGATGAGCGCATAGGCTTGGGCTG | 2340 |
| Qy | 2341 | GGCCGGGAGGTGGCGGTGTGGTCTGTGTCCCATCTTTCGCGCTTTGCCGTG | 2400 |
| Db | 2341 | GGCCGGGAGGTGGCGGTGTGGTCTGTGTCCCATCTTTCGCGCTTTGCCGTG | 2400 |
| Qy | 2401 | ATGACCGTGGCTATCTCTGCTGGTGAATGAGGAGGACTCTCAGCCCTTCTTGCAAGCCCTGCGG | 2460 |
| Db | 2401 | ATGACCGTGGCTATCTCTGCTGGTGAATGAGGAGGACTCTCAGCCCTTCTTGCAAGCCCTGCGG | 2460 |
| Qy | 2461 | CTGCATGGGTGGNAATTCAGAAACAGTTCTATCAGGACAGGGCTTCAAGCTGAGTCCC | 2520 |
| Db | 2461 | CTGCATGGGTGGNAATTCAGAAACAGTTCTATCAGGACAGGGCTTCAAGCTGAGTCCC | 2520 |
| Qy | 2521 | TTCACTTTCGCTGCCACAGATGACTAGGGGCCACTGCAAGTCTCTGCCAGACCTCTTCTCT | 2580 |
| Db | 2521 | TTCACTTTCGCTGCCACAGATGACTAGGGGCCACTGCAAGTCTCTGCCAGACCTCTTCTCT | 2580 |
| Qy | 2581 | GACCTCTGAGGAGGAGGAATAAGACGGTCCGCTCGCAAAAAAATAAAAAA | 2640 |
| Db | 2581 | GACCTCTGAGGAGGAGGAATAAGACGGTCCGCTCGCAAAAAAATAAAAAA | 2640 |

RESULT 9
ACA56496

ACA56496
ID ACA56496 standard; cDNA; 2655/BP.

ACA56496;

XX DT 06-JUN-2003 (first entry)

DE Human signalling pathway polynucleotide probe SEQ ID NO 1094.

Human: probe: ss: array element: Parkinson's disease:

KW human; probe; ss; array element; Parkinson's disease;
KW signalling pathway population: cancer; adenocarcinoma; leukaemia;

immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
Homo sapiens.
US6500938-B1.
31-DEC-2002.
30-JAN-1998; 98US-00016434.
30-JAN-1998; 98US-00016434.
(INCY-) INCYTE GENOMICS INC.
Au-Young J, Seilhamer JJ;
WPI; 2003-352189/33.
Combination of polynucleotide probes, useful as array elements in a
microarray for monitoring the expression of a number of target
polynucleotides.
Claim 1; SEQ ID NO 1094; 65pp; English.

Sequence 2655 BP: 460 A: 885 C: 824 G: 486 T: 0 U: 0 Other:

```
Query Match      100.0%; Score 2640; DB 10; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0;
```

| | | | |
|----|-----|---|-----|
| Qy | 1 | CGCGCTGGCGGACGGGAGCCACGACAGCGAGGCGCGGCGGACGACACACCGGGGACCATG | 60 |
| Db | 1 | CGGCGTGGCGGAGCGGGCAGCGACGACGAGCGAGCGCGCGGCGACACACCGGGGACCATG | 60 |
| Qy | 61 | GGCTCCATGTTCCGAGCGAGAGGTGGCCCTGTGTCCAGCTCTTTCTGCGCCACAGCGGGCT | 120 |
| Db | 61 | GGCTCCATGTTCCGAGCGAGAGGTGGCCCTGTGTCCAGCTCTTTCTGCGCCACAGCGGGCT | 120 |
| Qy | 121 | GCCTACACCTGCGTAGTCGGCTGGGCGAGCTGGGCGCTCGTGGAGTTCCAGAGACCTCAAC | 180 |
| Db | 121 | GCCTACACCTGCGTAGTCGGCTGGGCGAGCTGGGCGCTCGTGGAGTTCCAGAGACCTCAAC | 180 |
| Qy | 181 | GCCTCGGTAGCGCCTTCCAGAGACGCTTTGTGTTTGATTTTGGCGCTGTGAGGAGCTG | 240 |
| Db | 181 | GCCTCGGTAGCGCCTTCCAGAGACGCTTTGTGTTTGATTTTGGCGCTGTGAGGAGCTG | 240 |
| Qy | 241 | GAGAAGACCTTCACTTCTCCAGAGAGAGGTGCGCGGGCTGGGTGCTGCTGCCCCCG | 300 |
| Db | 241 | GAGAAGACCTTCACTTCTCCAGAGAGAGGTGCGCGGGCTGGGTGCTGCTGCCCCCG | 300 |

QY 301 CCAAGGGAGGCTGCGCAGACCCCAACCCCGGACCTGCTGCGATCCAGGAGAGCG 360
DB 301 CCAAGGGAGGCTGCGCAGACCCCAACCCCGGACCTGCTGCGATCCAGGAGAGCG 360
QY 361 GAGCGCTTGGCCAGGAGCTGCGGATGTGCGGGAJAAACAGCAGCCCTTGGCGCCAG 420
DB 361 GAGCGCTTGGCCAGGAGCTGCGGATGTGCGGGAJAAACAGCAGCCCTTGGCGCCAG 420
QY 421 CTGCAACAGCTGAGCTCCAGCCCGCTGCTACGCCAGGGCCATGAACCTCAGCTGGCA 480
DB 421 CTGCAACAGCTGAGCTCCAGCCCGCTGCTACGCCAGGGCCATGAACCTCAGCTGGCA 480
QY 481 GCCGCCACACAGATGGGCTCAGAGAGAGCCCTGCTCCAGGCCCCCGGGGGCGG 540
DB 481 GCCGCCACACAGATGGGCTCAGAGAGAGCCCTGCTCCAGGCCCCCGGGGGCGG 540
QY 541 CACCAAGACCTGAGGGTCAATTTGTGAGAGTGGCGTGGAGCCCAACAGGCCCTTGCC 600
DB 541 CACCAAGACCTGAGGGTCAATTTGTGAGAGTGGCGTGGAGCCCAACAGGCCCTTGCC 600
QY 601 CTAGAGCCCTGCTCTGAGGGCTGCGCGGCTTCTCATTTGCGAGCTTCAGGGAGCTG 660
DB 601 CTAGAGCCCTGCTCTGAGGGCTGCGCGGCTTCTCATTTGCGAGCTTCAGGGAGCTG 660
QY 661 GAGCAGCCGCTGAGCACCCTGTAGCGGCGAGCCAGCCAGCTGATGACCTTCTCATC 720
DB 661 GAGCAGCCGCTGAGCACCCTGTAGCGGCGAGCCAGCCAGCTGATGACCTTCTCATC 720
QY 721 TCCTACTGGGTGAGCAGATCGGAAGAGATCCGCAAGATCAACAGCTGCTTCACTGC 780
DB 721 TCCTACTGGGTGAGCAGATCGGAAGAGATCCGCAAGATCAACAGCTGCTTCACTGC 780
QY 781 CAGCTTCTCCGTTCTGAGCAGAGAGGCGCCCTGCGGGCTTCCAGCAGCTGCA 840
DB 781 CAGCTTCTCCGTTCTGAGCAGAGAGGCGCCCTGCGGGCTTCCAGCAGCTGCA 840
QY 841 CAGCAGCCAGGAGCTGAGAGGTCTCTCGGGAGAGCAGCGGTCTGAGCAGCTG 900
DB 841 CAGCAGCCAGGAGCTGAGAGGTCTCTCGGGAGAGCAGCGGTCTGAGCAGCTG 900
QY 901 CTAGCCCGGTGCTGAGCTGCTGCGCAGGCGAGGTGCAAGTCCACAGATGAAGCC 960
DB 901 CTAGCCCGGTGCTGAGCTGCTGCGCAGGCGAGGTGCAAGTCCACAGATGAAGCC 960
QY 961 GTGTACCTGGCCCTGAAACAGTGCAGCTGAGCAGCAGCTGCTCATTTGCCGAG 1020
DB 961 GTGTACCTGGCCCTGAAACAGTGCAGCTGAGCAGCAGCTGCTCATTTGCCGAG 1020
QY 1021 GCCTGGTCTCTGTGCGAGACCTGCGCCCTGCGAGAGGCGCTTCCGCGACAGCTCGATG 1080
DB 1021 GCCTGGTCTCTGTGCGAGACCTGCGCCCTGCGAGAGGCGCTTCCGCGACAGCTCGATG 1080
QY 1081 GAGGAGGAGTGTGCGCTGCTCACCGCATCCCTGCGGGAATGCGCCCACTC 1140
DB 1081 GAGGAGGAGTGTGCGCTGCTCACCGCATCCCTGCGGGAATGCGCCCACTC 1140
QY 1141 ATCCGCAACACCGCTTCAAGGCGAGCTTCAAGGCGATCGTGGATCGCTAGCGGTGGG 1200
DB 1141 ATCCGCAACACCGCTTCAAGGCGAGCTTCAAGGCGATCGTGGATCGCTAGCGGTGGG 1200
QY 1201 CGCTACAGGAGGTCAACCCCGCTTCTACACCATCATCCTTCCCTTCTGTTGCT 1260
DB 1201 CGCTACAGGAGGTCAACCCCGCTTCTACACCATCATCCTTCCCTTCTGTTGCT 1260
QY 1261 GTATGTTGGGGATGTGGGCAACCGGCTGCTCATGTTCTTCCCTTCCCTTCCCTTCCCT 1320
DB 1261 GTATGTTGGGGATGTGGGCAACCGGCTGCTCATGTTCTTCCCTTCCCTTCCCTTCCCT 1320
QY 1321 CTTGGGAGAACCGACCGCTGTGAAGCGCGAGAGAGAGTCTGCGACAGCTTCTTC 1380
DB 1321 CTTGGGAGAACCGACCGCTGTGAAGCGCGAGAGAGAGTCTGCGACAGCTTCTTC 1380
QY 1381 AGGGCCCGTACCTGCTCTGTTATGGGCGCTGTTCTCCATCTACACCGGCTTCATCTAC 1440

DB 1381 AGGGCCCGTACCTGCTCTGTTATGGGCGCTGTTCTCCATCTACACCGGCTTCATCTAC 1440
QY 1441 AACAGTGTCTTCAAGTGTGCGCCACACAGCATCTTCCCTCGGCTGAGTGTGCGCCCATG 1500
DB 1441 AACAGTGTCTTCAAGTGTGCGCCACACAGCATCTTCCCTCGGCTGAGTGTGCGCCCATG 1500
QY 1501 GCCAACAGTCTGCTGAGTGTGCAATCTTGGCCAGCAGCACAGATGCTTACCTGGAT 1560
DB 1501 GCCAACAGTCTGCTGAGTGTGCAATCTTGGCCAGCAGCACAGATGCTTACCTGGAT 1560
QY 1561 CCCAACGTCACCGGTGTCTTCTGGAGCCCTTACCCCTTGGCATCGATCTTATTTGGAGC 1620
DB 1561 CCCAACGTCACCGGTGTCTTCTGGAGCCCTTACCCCTTGGCATCGATCTTATTTGGAGC 1620
QY 1621 CTGGCTGCAACCACTTGAAGTCTTCAATCTTCAAGATGAAGATGTCGTCTATCTG 1680
DB 1621 CTGGCTGCAACCACTTGAAGTCTTCAATCTTCAAGATGAAGATGTCGTCTATCTG 1680
QY 1681 GGCGTGTGACATGCGCTTGGGGTGTCTCGAGTCTTCAACAGTGCACCTTTGGC 1740
DB 1681 GGCGTGTGACATGCGCTTGGGGTGTCTCGAGTCTTCAACAGTGCACCTTTGGC 1740
QY 1741 CAGAGGCAACCGCTGCTGCGAGACGCTGCGGAGTCACTTCTCTGCTGGGACTCTTTC 1800
DB 1741 CAGAGGCAACCGCTGCTGCGAGACGCTGCGGAGTCACTTCTCTGCTGGGACTCTTTC 1800
QY 1801 GGTACTCTGTTCTTAGTCACTTCAAGTGTGTGTCTGGCTGCGCAGGCGCGC 1860
DB 1801 GGTACTCTGTTCTTAGTCACTTCAAGTGTGTGTCTGGCTGCGCAGGCGCGC 1860
QY 1861 TCGCCAGCATCTCATCCACTTCAACATGTCTTCTTCTCCACAGCCCCAGCAAC 1920
DB 1861 TCGCCAGCATCTCATCCACTTCAACATGTCTTCTTCTCCACAGCCCCAGCAAC 1920
QY 1921 AGGTGCTCTTACCCCGGAGAGTGTTCAGGCGACGCTGCTGGTCTTGGCTTGGCC 1980
DB 1921 AGGTGCTCTTACCCCGGAGAGTGTTCAGGCGACGCTGCTGGTCTTGGCTTGGCC 1980
QY 1981 ATGTGCGCATCTGCTGCTTGGCACAACCTGCTGCACTGTGCGACCGCGCGC 2040
DB 1981 ATGTGCGCATCTGCTGCTTGGCACAACCTGCTGCACTGTGCGACCGCGCGC 2040
QY 2041 CTGCGAGGAGCGCGCTGACCGAGAGAAACAGGCGCGCTTCTGCACTGCT 2100
DB 2041 CTGCGAGGAGCGCGCTGACCGAGAGAAACAGGCGCGCTTCTGCACTGCT 2100
QY 2101 GACGATCTGTGAATGCTGAGCTTCCGATGAGGAAAGAGGCGGCTTGGATGATA 2160
DB 2101 GACGATCTGTGAATGCTGAGCTTCCGATGAGGAAAGAGGCGGCTTGGATGATA 2160
QY 2161 GAGAGGCGAGCTGCTCCCTCGAGTGTCTATGACAGGCGCATCCACACATCGAG 2220
DB 2161 GAGAGGCGAGCTGCTCCCTCGAGTGTCTATGACAGGCGCATCCACACATCGAG 2220
QY 2221 TTCTGCTGGGCTGCTGCTCAACACCGCTTCTTACCTGCGCTGCTGGGCTTGGGCTG 2280
DB 2221 TTCTGCTGGGCTGCTGCTCAACACCGCTTCTTACCTGCGCTGCTGGGCTTGGGCTG 2280
QY 2281 GCCACGCGCAGCTGCTGAGGTTCTGTTGGGCTTGTGTGATGCGCATAGGCTTGGGCTG 2340
DB 2281 GCCACGCGCAGCTGCTGAGGTTCTGTTGGGCTTGTGTGATGCGCATAGGCTTGGGCTG 2340
QY 2341 GCGCGGAGTGGGCTGCTGTTGGTGTGCTGCTTTCCTTTCGCGCTTTCGCGT 2400
DB 2341 GCGCGGAGTGGGCTGCTGTTGGTGTGCTGCTTTCCTTTCGCGCTTTCGCGT 2400
QY 2401 ATGACCGTGGCTATCTGCTGTGATGAGGAGTCTCAGCTTCTCAGCGCTTGG 2460
DB 2401 ATGACCGTGGCTATCTGCTGTGATGAGGAGTCTCAGCTTCTCAGCGCTTGG 2460
QY 2461 CTGCACTGGGTGGAATTCAGAAAGTTCTACTCAGGCGGCTACAAGCTGAGTCCC 2520

| | | |
|------|--|------|
| 901 | CTAGGCGGGTGTGTCGAGCTGCTGCCCGCAGGCGAGGTGCGAGGTTCACAAGATGAAGGCC | 960 |
| 961 | GTGTACTCTGGCCCTTGAACCAAGTGCAGCGTGTGAGCAACCAAGTGCCTCATTTGCCGAG | 1020 |
| 961 | GTGTACTCTGGCCCTTGAACCAAGTGCAGCGTGTGAGCAACCAAGTGCCTCATTTGCCGAG | 1020 |
| 1021 | GCCTGGTGTCTGTGTCGAGAACCTGCCCGCCTTGCAGAGAGGCCCTGCGGGAACAGTCTGATG | 1080 |
| 1021 | GCCTGGTGTCTGTGTCGAGAACCTGCCCGCCTTGCAGAGAGGCCCTGCGGGAACAGTCTGATG | 1080 |
| 1081 | GAGGAGGAGTGTAGTGTGGTGTCAACGCATCCCTCTCCGCGGACATGCCCCCAACACTC | 1140 |
| 1081 | GAGGAGGAGTGTAGTGTGGTGTCAACGCATCCCTCTCCGCGGACATGCCCCCAACACTC | 1140 |
| 1141 | ATTCGGCAACCAACCGCTTTCAGCGGCAGCTTCCAGGACATCTGTGTGATCTGCTACGAGTGGC | 1200 |
| 1141 | ATTCGGCAACCAACCGCTTTCAGCGGCAGCTTCCAGGACATCTGTGTGATCTGCTACGAGTGGC | 1200 |
| 1201 | CGCTACCAAGAGGTCAACCCCGCTCCCTACACATCATCACTTCCCTTCCCTTCTGTGTGCT | 1260 |
| 1201 | CGCTACCAAGAGGTCAACCCCGCTCCCTACACATCATCACTTCCCTTCCCTTCTGTGTGCT | 1260 |
| 1261 | GTGATGTTCCGGGATGTGGCCACAGGCTGTCTCATGTTTCTTCTTGGCCCTGGCCATGGTC | 1320 |
| 1261 | GTGATGTTCCGGGATGTGGCCACAGGCTGTCTCATGTTTCTTCTTGGCCCTGGCCATGGTC | 1320 |
| 1321 | CTTTCGGAGAACCGACCGGTGTGAAGCCGCGAGAACGAGATCTGGCAGACTTCTTCTTC | 1380 |
| 1321 | CTTTCGGAGAACCGACCGGTGTGAAGCCGCGAGAACGAGATCTGGCAGACTTCTTCTTCTTC | 1380 |
| 1381 | AGGGGCGCTACCTGTCTCTGTCTTATGGGCTGTGTTCTTCATCTACACCGGCTTCATCTAC | 1440 |
| 1381 | AGGGGCGCTACCTGTCTCTGTCTTATGGGCTGTGTTCTTCATCTACACCGGCTTCATCTAC | 1440 |
| 1441 | AAAGAGTGTCTTCTAGTTCGGGCAACAGGATCTTCCCTCTGGGCTGTGAGTGTGGCGGCCATG | 1500 |
| 1441 | AAAGAGTGTCTTCTAGTTCGGGCAACAGGATCTTCCCTCTGGGCTGTGAGTGTGGCGGCCATG | 1500 |
| 1501 | GCCAAACAGTCTGGCTGGAGTGTGATCTTCTGGCCAGCAGACGATGCTTACCTTGGAT | 1560 |
| 1501 | GCCAAACAGTCTGGCTGGAGTGTGATCTTCTGGCCAGCAGACGATGCTTACCTTGGAT | 1560 |
| 1561 | CCCAAACGTCAACGGTGTCTTCTGGGACCTTAACCCCTTGGCATTCGATCTTATTTGGAGC | 1620 |
| 1561 | CCCAAACGTCAACGGTGTCTTCTGGGACCTTAACCCCTTGGCATTCGATCTTATTTGGAGC | 1620 |
| 1621 | CTGGCTGCCAAACACTTGTGCTTCTCAACTCTTCAAGATGAGATGTCCGTCATCTCTG | 1680 |
| 1621 | CTGGCTGCCAAACACTTGTGCTTCTCAACTCTTCAAGATGAGATGTCCGTCATCTCTG | 1680 |
| 1681 | GGCGTGTGACATGCGCTTGTGGGTTGGTCTCTGGAGTCTTCAACACGCTGCACCTTTGGC | 1740 |
| 1681 | GGCGTGTGACATGCGCTTGTGGGTTGGTCTCTGGAGTCTTCAACACGCTGCACCTTTGGC | 1740 |
| 1741 | CAGAGGACCGGCTGTCTGTGGAGACGCTGCGGAGCTCACTTCTCTGTCTGGAGACTCTTTC | 1800 |
| 1741 | CAGAGGACCGGCTGTCTGTGGAGACGCTGCGGAGCTCACTTCTCTGTCTGGAGACTCTTTC | 1800 |
| 1801 | GGTTACTCTCTGTGTTCTTGTGATCTCAAGTGGCTGTGCTGTGGGCTGCCAGGGCGGCC | 1860 |
| 1801 | GGTTACTCTCTGTGTTCTTGTGATCTCAAGTGGCTGTGCTGTGGGCTGCCAGGGCGGCC | 1860 |
| 1861 | TGCGCCAGCATCTCTATCCACTTCAATGTTTCTTCTTCTTCCCAACAGCCCCCAACAAC | 1920 |
| 1861 | TGCGCCAGCATCTCTATCCACTTCAATGTTTCTTCTTCTTCCCAACAGCCCCCAACAAC | 1920 |
| 1921 | AGGCTGTCTTACCCCGGAGGAGGTGGTCCAGGCCACGCTGTGTGCTCTGGCTTGGCC | 1980 |
| 1921 | AGGCTGTCTTACCCCGGAGGAGGTGGTCCAGGCCACGCTGTGTGCTCTGGCTTGGCC | 1980 |
| 1981 | ATGGTGCCCATCTCTGTCTGTGGCACCCTCTGCACCTCTGTGACGCGCACCGCCGCGCG | 2040 |
| 1981 | ATGGTGCCCATCTCTGTCTGTGGCACCCTCTGTGACGCGCACCGCCGCGCACCGCCGCGCG | 2040 |

| | | | | |
|----|------|---------------------------------------|--------------------------------------|------|
| Qy | 2041 | CTGCGGAGGAGCGCCGTGACCGACAGAGGAAAAA | CAAGGCCGGGTGCTGGACCTGCGCT | 2100 |
| Db | 2041 | CTGCGGAGGAGCGCCGTGACCGACAGGAGAAAAA | CAAGGCCGGGTGCTGGACCTGCGCT | 2100 |
| Qy | 2101 | GAGCATCTGTGAATGCTCGAGCTCCGATGAGGAAAAG | CAGGGGCGCTGGATGATGAA | 2160 |
| Db | 2101 | GAGCATCTGTGAATGCTCGAGCTCCGATGAGGAAAAG | CAGGGGCGCTGGATGATGAA | 2160 |
| Qy | 2161 | GAGGAGCCGAGCTCGTCCCTCCGAGGTGCTCATG | CACAGGCCCATCCACACCATCGAG | 2220 |
| Db | 2161 | GAGGAGCCGAGCTCGTCCCTCCGAGGTGCTCATG | CACAGGCCCATCCACACCATCGAG | 2220 |
| Qy | 2221 | TTCTGCTCGGGCTGCGTCTCCAA | CACCGCTCTTACCTGCGCCTGTGGGCCCTGAGCCTG | 2280 |
| Db | 2221 | TTCTGCTCGGGCTGCGTCTCCAA | CACCGCTCTTACCTGCGCCTGTGGGCCCTGAGCCTG | 2280 |
| Qy | 2281 | GCCACGCCAGCTGTCCGAGGTTCTGTGGGCCCAT | GGTGATGCGCATAGGCCCTGGGGCCTG | 2340 |
| Db | 2281 | GCCACGCCAGCTGTCCGAGGTTCTGTGGGCCCAT | GGTGATGCGCATAGGCCCTGGGGCCTG | 2340 |
| Qy | 2341 | GGCCGGGAGGTGGGCGCTGTGTGGTCTCGTCCCAT | CTTTGTCGGCTTTTGCGCTG | 2400 |
| Db | 2341 | GGCCGGGAGGTGGGCGCTGTGTGGTCTCGTCCCAT | CTTTGTCGGCTTTTGCGCTG | 2400 |
| Qy | 2401 | ATGACCTGTGCTATCTCTGTGTGTGATGAGGGAG | CTCTAGCCTTCCTGTGACGCCCTGCGG | 2460 |
| Db | 2401 | ATGACCTGTGCTATCTCTGTGTGTGATGAGGGAG | CTCTAGCCTTCCTGTGACGCCCTGCGG | 2460 |
| Qy | 2461 | CTGCACCTGGGTGGAAATCCAGAACAGTTCTACT | CTCAGGCCACGGGCTTACAGCTGAGTCCC | 2520 |
| Db | 2461 | CTGCACCTGGGTGGAAATCCAGAACAGTTCTACT | CTCAGGCCACGGGCTTACAGCTGAGTCCC | 2520 |
| Qy | 2521 | TTCACTTTCGCTGCCACAGATGACTAGGGGCCACT | TCGAGGTCCCTGCCAGACCTCCTTCCT | 2580 |
| Db | 2521 | TTCACTTTCGCTGCCACAGATGACTAGGGGCCACT | TCGAGGTCCCTGCCAGACCTCCTTCCT | 2580 |
| Qy | 2581 | GACCTCTGAGGCCAGGAGGAAATAAGACGGTCCG | CCCTGGCAAAAAA | 2640 |
| Db | 2581 | GACCTCTGAGGCCAGGAGGAAATAAGACGGTCCG | CCCTGGCAAAAAA | 2640 |

RESULT 11

ADT56292

ADI56292
ID ADI56292 standard: DNA: 2655 BP.

ID ADI56292
XX
AC ADI56292;

DT /22-APR-2004 (first entry)

XX Human polynucleotide probe #1094.

Human; probe; ss; receptor-like polypeptide; transducing polypeptide; effector-like polypeptide; cancer; immunopathology; drug development; toxicology; carcinogenicity; signalling pathway polypeptide; adrenal gland; bone marrow; brain; breast; cervix; tumour; immunopathology; AIDS; diabetes; pancreatitis; osteoporosis; ulcerative colitis; neuropathology; dementia; amnesia; epilepsy; Alzheimer's disease; depression.

OS Homo sapiens.

XX PN US2004010136-A1.

PD 15-JAN-2004:

26-NOV-2002; 2002US-00305720.

XX
PR 30-JAN-1998; 98US-00016434.

XX
PA (INCY-) INCYTE GENOMICS INC.

XX
PI Au-Young J, Seilhamer JJ;

| | | | |
|----|------|---|------|
| Db | 1561 | CCCAACGTCACCGGTGTCCTTCTCTGGGACCCCTACCCCTTGGCATCGATCTCTATTGGAGC | 1620 |
| Qy | 1621 | CTGCGCTGCACCAACTTGAGCTTCCTCAACTCCTTTCAAGATGAAGATGTCGGTCACTCCTG | 1680 |
| Db | 1621 | CTGCGCTGCCAACCACTTGAGCTTCCTCAACTCCTTTCAAGATGAAGATGTCGGTCACTCCTG | 1680 |
| Qy | 1681 | GGCGTCGTGCACATGGCGCTTTGGGGTGGTCCCTCGGAGTCTTTCAACACAGTGCATCTTGGC | 1740 |
| Db | 1681 | GGCGTCGTGCACATGGCGCTTTGGGGTGGTCCCTCGGAGTCTTTCAACACAGTGCATCTTGGC | 1740 |
| Qy | 1741 | CAGAGGCACCGGCTGCTGTCTGGAGACCGTGC CGGAGCTCACTTCTCTGTCTGGAGCTTCTT | 1800 |
| Db | 1741 | CAGAGGCACCGGCTGCTGTCTGGAGACCGTGC CGGAGCTCACTTCTCTGTCTGGAGCTTCTT | 1800 |
| Qy | 1801 | GGTTACTCTCGTGTTCCTAGTTCATCTACAAGTGGCTGTGTCTGGGCTGCACAGGCGCCG | 1860 |
| Db | 1801 | GGTTACTCTCGTGTTCCTAGTTCATCTACAAGTGGCTGTGTCTGGGCTGCACAGGCGCCG | 1860 |
| Qy | 1861 | TCGCCACAGCATCTTCATCCACTTCATCAACAATGTTCTCTTCTCCACAGCCCCAGCAAC | 1920 |
| Db | 1861 | TCGCCACAGCATCTTCATCCACTTCATCAACAATGTTCTCTTCTCCACAGCCCCAGCAAC | 1920 |
| Qy | 1921 | AGGCTGCTCTACCCCCCGGACAGGAGTGGTCCAGGCCACGCTGTGTGTCTGGCTTGGCC | 1980 |
| Db | 1921 | AGGCTGCTCTACCCCCCGGACAGGAGTGGTCCAGGCCACGCTGTGTGTCTGGCTTGGCC | 1980 |
| Qy | 1981 | ATGCTGCCCATCTCTGCTGTGGGCACACCCCTGCACCTGTGCACCGCCACCGCGCGCCG | 2040 |
| Db | 1981 | ATGCTGCCCATCTCTGCTGTGGGCACACCCCTGCACCTGTGCACCGCCACCGCGCGCCG | 2040 |
| Qy | 2041 | CTCGGAGGAGCGCCGTGCACCGACAGGAGAAAACAGGCCGGTGTCTGGACCTGCGCT | 2100 |
| Db | 2041 | CTCGGAGGAGCGCCGTGCACCGACAGGAGAAAACAGGCCGGTGTCTGGACCTGCGCT | 2100 |
| Qy | 2101 | GACGCATCTGTGAATGCTGGAGCTCCGATGAGGAAAAAGCAGGGGCCCTGGATGATGAA | 2160 |
| Db | 2101 | GACGCATCTGTGAATGCTGGAGCTCCGATGAGGAAAAAGCAGGGGCCCTGGATGATGAA | 2160 |
| Qy | 2161 | GAGGAGCGCAGCTCGTCCCTCCGAGGTGCTCATGCACACGAGGCATCCACACCATCGAG | 2220 |
| Db | 2161 | GAGGAGCGCAGCTCGTCCCTCCGAGGTGCTCATGCACACGAGGCATCCACACCATCGAG | 2220 |
| Qy | 2221 | TTCTGCTGGGCTGCGTCTCCAACACCGCTCTTACTGCGCCTGTGGGCCCTTGAGCCTG | 2280 |
| Db | 2221 | TTCTGCTGGGCTGCGTCTCCAACACCGCTCTTACTGCGCCTGTGGGCCCTTGAGCCTG | 2280 |
| Qy | 2281 | GCCACGCCAGCTGTCGAGGTCTGTGGGCCATGTGTGATGCGCATAGGCCCTGGGCCCTG | 2340 |
| Db | 2281 | GCCACGCCAGCTGTCGAGGTCTGTGGGCCATGTGTGATGCGCATAGGCCCTGGGCCCTG | 2340 |
| Qy | 2341 | GGCCGGAGGTGGGCGTGGGGGCTGTGGTGTGGTCCCCATCTTTGGCGCCTTTTGGCGTG | 2400 |
| Db | 2341 | GGCCGGAGGTGGGCGTGGGGGCTGTGGTGTGGTCCCCATCTTTGGCGCCTTTTGGCGTG | 2400 |
| Qy | 2401 | ATGACCGTGGCTATCTCTGTGTGTGATGGAAGGACTCTACGCGTTCCTGCACGCGCTGCGG | 2460 |
| Db | 2401 | ATGACCGTGGCTATCTCTGTGTGTGATGGAAGGACTCTCTCAGCCCTTCTCTGCACGCGCTGCGG | 2460 |
| Qy | 2461 | CTGCACCTGGGTGGAATTCACAGAACAGTCTTACTCAGGCAAGGCTTACAGCTGAGTCCC | 2520 |
| Db | 2461 | CTGCACCTGGGTGGAATTCACAGAACAGTCTTACTCAGGCAAGGCTTACAGCTGAGTCCC | 2520 |
| Qy | 2521 | TTACCTTTCGTCGACAGATGACTAGGGGCCACTGCAGGTCTCTGCCAGACCTCTTCTCT | 2580 |
| Db | 2521 | TTACCTTTCGTCGACAGATGACTAGGGGCCACTGCAGGTCTCTGCCAGACCTCTTCTCT | 2580 |
| Qy | 2581 | GACCTCTGAGCGAGGAGGAAATAAGACGGTTCGCCCTGGCAAAAAAATAAAAAA | 2640 |
| Db | 2581 | GACCTCTGAGCGAGGAGGAAATAAGACGGTTCGCCCTGGCAAAAAAATAAAAAA | 2640 |

| | | |
|----|---|--|
| AD | AD833657 | |
| ID | AD833657 standard; cDNA; 2655 BP. | |
| XX | | |
| AC | AD833657; | |
| XX | | |
| DT | 11-AUG-2005 (first entry) | |
| XX | | |
| DE | Human lymph node cDNA #916. | |
| XX | | |
| KW | ss; gene; human; immunological response; blood cell; cancer; | |
| KW | immunopathological; AIDS; allergy; anaemia; asthma; atherosclerosis; | |
| KW | bronchitis; ulcerative colitis; diabetes; multiple sclerosis; | |
| KW | osteoporosis; pancreatitis; infection; arthritis; lymph node. | |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| PN | US2004077003-A1. | |
| XX | | |
| PD | 22-APR-2004. | |
| XX | | |
| PF | 14-AUG-2003; 2003US-00641643. | |
| XX | | |
| PR | 09-FEB-1998; 98US-00023655. | |
| XX | | |
| PA | (INCY-) INCYTE CORP. | |
| XX | | |
| PI | Cocks BG, Stuart SG, Seilhamer JJ; | |
| XX | | |
| DR | WPI; 2004-387937/36. | |
| XX | | |
| PT | New compositions having a number of first, second and third | |
| PT | polynucleotide probes, useful in research and diagnostic applications in | |
| PT | cancer and immunological conditions e.g. AIDS, diabetes, osteoporosis and | |
| PT | infections. | |
| XX | | |
| PS | Claim 15; SEQ ID NO 916; 16pp; English. | |
| XX | | |
| CC | The invention relates to polynucleotides which are used as probes to | |
| CC | detect genes differentially expressed in an immunological response, | |
| CC | abundantly expressed in an immunological response and/or coding for a | |
| CC | polypeptide known to regulate blood cell biology. The polynucleotides are | |
| CC | useful in research and diagnostic applications particularly in cancer and | |
| CC | immunopathological conditions, such as AIDS, allergies, anaemia, asthma, | |
| CC | atherosclerosis, bronchitis, ulcerative colitis, diabetes, multiple | |
| CC | sclerosis, osteoporosis, pancreatitis, infections and arthritis. The | |
| CC | present sequence represents a human lymph node cDNA used to detect blood | |
| CC | cell and immunological response gene expression. Note: The present | |
| CC | sequence does not appear in the printed specification but was obtained in | |
| CC | electronic format from the USPTO web site | |
| CC | (seqdata.uspto.gov/sequence.html?DocID=20040077003). | |

QY 241 GAGAGACCTTACCTTCTCCAGGAGGAGGTGCGCGGCTGGGTCTGTCGCCCGC 300
DB 241 GAGAGACCTTACCTTCTCCAGGAGGAGGTGCGCGGCTGGGTCTGTCGCCCGC 300
QY 301 CCAAGGGGAGGTGCGCGCACCCACCCCGGAGCCTGTGCGCATCCAGGAGGAGACG 360
DB 301 CCAAGGGGAGGTGCGCGCACCCACCCCGGAGCCTGTGCGCATCCAGGAGGAGACG 360
QY 361 GAGGCGCTGCGCCAGGAGCTGCGGGATGTGCGGGGCAACACAGAGGCCCTGCGGGCCAG 420
DB 361 GAGGCGCTGCGCCAGGAGCTGCGGGATGTGCGGGGCAACACAGAGGCCCTGCGGGCCAG 420
QY 421 CTCACACAGCTGACAGCTCCAGCGCGCTGCTAGCGCAGGCGCATGAACCTCAGCTGGCA 480
DB 421 CTCACACAGCTGACAGCTCCAGCGCGCTGCTAGCGCAGGCGCATGAACCTCAGCTGGCA 480
QY 481 GCCGCCACACAGATGGGGCTCAGAGAGGACGCCCTGTCTCAGGCGCCCGGGGGCGG 540
DB 481 GCCGCCACACAGATGGGGCTCAGAGAGGACGCCCTGTCTCAGGCGCCCGGGGGCGG 540
QY 541 CACAGGAACCTGAGGGTCAACTTTGTGGCAGGTGCCGTGAGGCCCAACAGGCCCTTGCC 600
DB 541 CACAGGACCTGAGGGTCAACTTTGTGGCAGGTGCCGTGAGGCCCAACAGGCCCTTGCC 600
QY 601 CTAGAGGCGCTGCTGAGGGGCTGCGCGGCTTCTCATTTGCGAGCTTCAGGAGCTG 660
DB 601 CTAGAGGCGCTGCTGAGGGGCTGCGCGGCTTCTCATTTGCGAGCTTCAGGAGCTG 660
QY 661 GAGCAGCGCTGAGACACCCCGTGAGCGGCGAGCCAGCCAGCTGGATGACCTTCTCATC 720
DB 661 GAGCAGCGCTGAGACACCCCGTGAGCGGCGAGCCAGCCAGCTGGATGACCTTCTCATC 720
QY 721 TCTACTGCGGTGAGCAGATCGGACAGAGATCCGCAAGATCAAGCTGCTTCCACTGC 780
DB 721 TCTACTGCGGTGAGCAGATCGGACAGAGATCCGCAAGATCAAGCTGCTTCCACTGC 780
QY 781 CAGCTCTTCCGTTCTGACAGGAGGAGGCCCGGCTCGGGCCCTGACAGCTGCA 840
DB 781 CAGCTCTTCCGTTCTGACAGGAGGAGGCCCGGCTCGGGCCCTGACAGCTGCA 840
QY 841 CAGCAGGCGAGGCTGACAGAGGTCTCGGGGAGACAGAGCGTTCTTGAGCCAGGTG 900
DB 841 CAGCAGGCGAGGCTGACAGAGGTCTCGGGGAGACAGAGCGTTCTTGAGCCAGGTG 900
QY 901 CTAGGCGGGTGTGAGCTGTGCGCGCAGGAGGAGTGCAGTCCAAAGATGAAGGCC 960
DB 901 CTAGGCGGGTGTGAGCTGTGCGCGCAGGAGGAGTGCAGTCCAAAGATGAAGGCC 960
QY 961 GTGTACTGCGCTGAACCAAGTGCAGGTGAGCACACAGCAAGTGCCTATTGCGGAG 1020
DB 961 GTGTACTGCGCTGAACCAAGTGCAGGTGAGCACACAGCAAGTGCCTATTGCGGAG 1020
QY 1021 GCCTGTGTCTGTGAGACCTGCGCGCTCGAGAGGCCCTGCGGGACAGCTCGATG 1080
DB 1021 GCCTGTGTCTGTGAGACCTGCGCGCTCGAGAGGCCCTGCGGGACAGCTCGATG 1080
QY 1081 GAGGAGGAGTGAAGTGCCTGAGTGCAGGATCCCTGCGGGACATGCCGCCACACTC 1140
DB 1081 GAGGAGGAGTGAAGTGCCTGAGTGCAGGATCCCTGCGGGACATGCCGCCACACTC 1140
QY 1141 ATCCGACCAACCGCTTACGGCCAGCTTCAGGGGATCGTGGATCGTACGGCGTGGG 1200
DB 1141 ATCCGACCAACCGCTTACGGCCAGCTTCAGGGGATCGTGGATCGTACGGCGTGGG 1200
QY 1201 CGCTACAGAGGTCAACCCCGCTCCCTACACCATCATACCTTCCCTTCTGTTGCT 1260
DB 1201 CGCTACAGAGGTCAACCCCGCTCCCTACACCATCATACCTTCCCTTCTGTTGCT 1260
QY 1261 GTGATGTTGGGAGTGGGCCACCGGCTGCTCATGTTCTTCTGCGCCCTGGCCATGGTC 1320
DB 1261 GTGATGTTGGGAGTGGGCCACCGGCTGCTCATGTTCTTCTGCGCCCTGGCCATGGTC 1320
QY 1321 CTTGGGAGAACCGACCGGCTGTGAAGCGCGCAGAACGAGATCTGGCAGACTTTCTTC 1380

DB 1321 CTTGGGAGAACCGACCGGCTGTGAAGCGCGCAGAACGAGATCTGGCAGACTTTCTTC 1380
QY 1381 AGGGCGCGCTACCTGTCTGCTTATGGGCTGTCTTCCATCTACACCGCTTCATCTAC 1440
DB 1381 AGGGCGCGCTACCTGTCTGCTTATGGGCTGTCTTCCATCTACACCGGCTTCATCTAC 1440
QY 1441 AACGAGTGTTCAGTCCGCGCACAGCATCTTCCCCTCGGGCTGGAGTGGCGCCCATG 1500
DB 1441 AACGAGTGTTCAGTCCGCGCACAGCATCTTCCCCTCGGGCTGGAGTGGCGCCCATG 1500
QY 1501 GCCAACAGTCTGGCTGGAGTGCATTCCTGGCCACACAGTGCATTCACCTGGAT 1560
DB 1501 GCCAACAGTCTGGCTGGAGTGCATTCCTGGCCACACAGTGCATTCACCTGGAT 1560
QY 1561 CCCAACGTCACCGGTGTCTTCTGGGACCTACCCCTTTGGCATCGATCTATTGGAGC 1620
DB 1561 CCCAACGTCACCGGTGTCTTCTGGGACCTACCCCTTTGGCATCGATCTATTGGAGC 1620
QY 1621 CTGGCTGCCAACCACTTGAGCTTCTCAACTCTTCAAGATGAAGATGTCGTCATCTG 1680
DB 1621 CTGGCTGCCAACCACTTGAGCTTCTCAACTCTTCAAGATGAAGATGTCGTCATCTG 1680
QY 1681 GCGCTGTCACATGCGCTTTGGGGTGTCTCGGAGTCTTCAACACAGTGCATTTGGC 1740
DB 1681 GCGCTGTCACATGCGCTTTGGGGTGTCTCGGAGTCTTCAACACAGTGCATTTGGC 1740
QY 1741 CAGAGCACCGGCTGTCTGGAGACGCTGCCGGAGCTCACTTCTCTGCTGGGACTCTTC 1800
DB 1741 CAGAGCACCGGCTGTCTGGAGACGCTGCCGGAGCTCACTTCTCTGCTGGGACTCTTC 1800
QY 1801 GGTACTCTGTTTCTAGTCACTAAGTGGCTGTGTCTGGGCTGCAGGGCGGCC 1860
DB 1801 GGTACTCTGTTTCTAGTCACTAAGTGGCTGTGTCTGGGCTGCAGGGCGGCC 1860
QY 1861 TCGCCAGCATCTCATCCACTTTCATCAACATGTTCTCTTCTCCACAGCCCGCAGCAAC 1920
DB 1861 TCGCCAGCATCTCATCCACTTTCATCAACATGTTCTCTTCTCCACAGCCCGCAGCAAC 1920
QY 1921 AGGCTCTCTACCCCCGCGAGGAGTGTCCAGGCCACGCTGTGTCTTCTGGCTTGGCC 1980
DB 1921 AGGCTCTCTACCCCCGCGAGGAGTGTTCAGGCCACGCTGTGTCTTCTGGCTTGGCC 1980
QY 1981 ATGTTGCCCATCTGTCTGTGTCACACCCCTGCACTGTGTCACCGCCACCGCCGCGC 2040
DB 1981 ATGTTGCCCATCTGTCTGTGTCACACCCCTGCACTGTGTCACCGCCACCGCCGCGC 2040
QY 2041 CTGCGGAGGAGGCCCTGACCGACAGGAGGAGAACAGGCCGCGGTTGCTGGACCTGCCT 2100
DB 2041 CTGCGGAGGAGGCCCTGACCGACAGGAGGAGAACAGGCCGCGGTTGCTGGACCTGCCT 2100
QY 2101 GAGCATCTGTGAATGGCTGGAGTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
DB 2101 GAGCATCTGTGAATGGCTGGAGTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
QY 2161 GAGGAGGCGAGCTGCTCCCTCCGAGGTCTCATGCAACAGGCCATCCACACCATCGAG 2220
DB 2161 GAGGAGGCGAGCTGCTCCCTCCGAGGTCTCATGCAACAGGCCATCCACACCATCGAG 2220
QY 2221 TTTCTGCTGGGCTGCTTCCAAACACCGCTCTACTTGGGCTGTGGGCTGTGAGGCTG 2280
DB 2221 TTTCTGCTGGGCTGCTTCCAAACACCGCTCTACTTGGGCTGTGGGCTGTGAGGCTG 2280
QY 2281 GCCACGCGCAGCTGCTCGAGGTTCTGTGGGCTGTGTGATGCGCATAGGCTTGGGCTG 2340
DB 2281 GCCACGCGCAGCTGCTCGAGGTTCTGTGGGCTGTGTGATGCGCATAGGCTTGGGCTG 2340
QY 2341 GCGCGGAGGTGGGCTGGGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
DB 2341 GCGCGGAGGTGGGCTGGGCTGGGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
QY 2401 ATGACCGTGGCTATCTGCTGTGTGATGGAGGAGTCTCAGCCTTCTCAGCGCCCTGCGG 2460

Dd 2401 ATGACCGTGGCTATCTGCTGGTGATGGAGGACTCTCAGCCCTCTCTGACGCCCTGCGG 2460
Qy 2461 CTGCACTGGGTGGATTCAGAACAGTTCTACTCAGGCAACGGCTACAGCTGAGTCCC 2520
Dd 2461 CTGCACTGGGTGGATTCAGAACAGTTCTACTCAGGCAACGGCTACAGCTGAGTCCC 2520
Qy 2521 TTACACCTTCGCTGCCACAGATGACTAGGCGCCCACTGACGGTCTCTGACAGACCTCTCTTCT 2580
Dd 2521 TTACACCTTCGCTGCCACAGATGACTAGGCGCCCACTGACGGTCTCTGACAGACCTCTCTTCT 2580
Qy 2581 GACCTCTGAGGAGGAGAGAGAAATAAGACGTCGCGCCCTGGCAAAAAA 2640
Dd 2581 GACCTCTGAGGAGGAGAGAGAAATAAGACGTCGCGCCCTGGCAAAAAA 2640

RESULT 13

AD128819
ID AD128819 standard; cdNA; 2700 BP.

AC AD128819;

XX 22-APR-2004 (first entry)

XX Human modifier of p53 (MP53) gene TCIRG-1.

XX Human; MP53; modifier of p53; p53; cytosolic; gene therapy;

KW T-cell immune regulator 1; TCIRG-1; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FH 109..2601
CDs /tag= a
FT /product= "MP53 TCIRG-1"

XX WP004004766-A1.
XX 15-JAN-2004.

XX 09-JUL-2003; 2003WO-US021378.
XX 10-JUL-2002; 2002US-0394992P.
XX 07-AUG-2002; 2002US-0401604P.
XX 16-SEP-2002; 2002US-0410988P.
XX 25-NOV-2002; 2002US-0428837P.

XX (EXEL-) EXELIXIS INC.
XX Costa MA, Maxwell ME, Lackner MR, Hung T, O'Brien CL, Jin Y;
XX Nicoll M, Hai B, Zhang H, Lickteig K, Amundsen CD;
XX WPI: 2004-142922/14.
XX P-PSDB; AD128834.
XX RFSFQ; NM_006019.2.

XX Identifying a candidate p53 pathway modulating agent for treating e.g.,
XX cancer, comprises contacting an assay system comprising a MP53
XX polypeptide or nucleic acid with a test agent and detecting a test agent-
XX biased activity.
XX Example 2; SEQ ID NO 10; 139pp; English.

XX In the present invention, genetic screens were designed to identify
XX modifiers of the p53 pathway in caenorhabditis elegans, where a
XX homozygous p53 deletion mutant was used. Various specific genes were
XX silenced by RNA inhibition. Genes causing altered phenotypes in the worms
XX were identified as modifiers of the p53 pathway. Human orthologs, denoted
XX modifiers of p52 (MP53), of these modifiers were identified. These
XX include the present sequence, characterised as encoding R-cell immune
XX regulator 1, ATPase, H+ transporting, or lysosomal V0 protein a isoform
XX 3. MP53 nucleic acids and polypeptides are attractive drug targets for
XX the treatment of pathologies associated with a defective p53 signalling
XX pathway, such as cancer. Methods for modulating MP53 function and/or the

CC p53 pathway in a mammalian cell involve contacting the cell with an agent
CC that specifically binds a MP53 polypeptide or nucleic acid. The agent may
CC be a small molecular modulator, a nucleic acid modulator or an antibody.
XX Sequence 2700 BP; 451 A; 905 C; 850 G; 494 T; 0 U; 0 Other;

Query Match 98.2%; Score 2592.2; DB 12; Length 2700;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2599; Conservative 0; Mismatches 9; Indels 3; Gaps 1;

Qy 33 GCGCGCGCAGCACACCCGCGGACCATGGGCTTCATGTTCCGAGAGAGAGTGGCCCT 92
Dd 84 GCGCGCGCAGCACACCCGCGGACCATGGGCTTCATGTTCCGAGAGAGAGTGGCCCT 143
Qy 93 GGTCCAGCTCTTTCTGCCCCACAGCGGCTGCTACACCTGCGTGAGTGGCTGGGCGAGCT 152
Dd 144 GGTCCAGCTCTTTCTGCCCCACAGCGGCTGCTACACCTGCGTGAGTGGCTGGGCGAGCT 203
Qy 153 GGGCCTCGTGGAGTTTCAGAGACCTCAACGCTCGGTGAGCGCTTCCAGAGACGCTTTCT 212
Dd 204 GGGCCTCGTGGAGTTTCAGAGACCTCAACGCTCGGTGAGCGCTTCCAGAGACGCTTTCT 263
Qy 213 GGTTCATGTTGGCGCTGTGAGGAGCTGGAGAGACCTTTCACCTTCTCTGACAGAGAGT 272
Dd 264 GGTTCATGTTGGCGCTGTGAGGAGCTGGAGAGACCTTTCACCTTCTCTGACAGAGAGT 323
Qy 273 GCGCGCGCTGGGCTGGTCTGCCCCCGCCAAAAGGGAGGCTGCGGCGACACCCCGG 332
Dd 324 GCGCGCGCTGGGCTGGTCTGCCCCCGCCAAAAGGGAGGCTGCGGCGACACCCCGG 383
Qy 333 GGCCTGCTGCGCATCCAGGAGGAGAGCGGCTGCGCCAGGAGAGCTGCGGAGTGGG 392
Dd 384 GGCCTGCTGCGCATCCAGGAGGAGAGCGGCTGCGCCAGGAGAGCTGCGGAGTGGG 443
Qy 393 GGGCAACAGCAGAGCGCTGCGGCGCCAGCTGCACAGCTGCAGCTCCACGCGCGCTGCT 452
Dd 444 GGGCAACAGCAGAGCGCTGCGGCGCCAGCTGCACAGCTGCAGCTCCACGCGCGCTGCT 503
Qy 453 AGCCAGCGGCGCATGAACCTCAGCTGGCAGCGCGCCACACAGATGGGGCTTCAGAGAGAC 512
Dd 504 AGCCAGCGGCGCATGAACCTCAGCTGGCAGCGCGCCACACAGATGGGGCTTCAGAGAGAC 563
Qy 513 GCGCCTGCTCAGAGCGCGCGGGGCGCCACACAGAGACCTGAGGGTCAACTTTGTGCGAG 572
Dd 564 GCGCCTGCTCAGAGCGCGCGGGGCGCCACACAGAGACCTGAGGGTCAACTTTGTGCGAG 623
Qy 573 TGCGGTGAGGCGCCACAAAGGCGCTGCGCTAGAGCGCTGCTCTGAGAGGCTGCGCGCG 632
Dd 624 TGCGGTGAGGCGCCACAAAGGCGCTGCGCTAGAGCGCTGCTCTGAGAGGCTGCGCGCG 683
Qy 633 CTTCCTCATTTGCCAGCTTCAGGAGCTGGAGCAGCGGCTGGAGCAGCGCGCTGACCGGCGA 692
Dd 684 CTTCCTCATTTGCCAGCTTCAGGAGCTGGAGCAGCGGCTGGAGCAGCGCGCTGACCGGCGA 743
Qy 693 GCGCAGCAGCTGGAGTCACTTCTCATCTCTGAGGCTGAGCAGATCGGACAGAGAT 752
Dd 744 GCGCAGCAGCTGGAGTCACTTCTCATCTCTGAGGCTGAGCAGATCGGACAGAGAT 803
Qy 753 CGCAAGATCAGGACTGCTTCCACTGCCACGCTCTTCCCGTTTCTGACAGAGAGAGGC 812
Dd 804 CGCAAGATCAGGACTGCTTCCACTGCCACGCTCTTCCCGTTTCTGACAGAGAGAGGC 863
Qy 813 CGGCTTCGGGCGCTTCAGCAGCTGCAACAGCAGAGCGGCTGAGCAGCGCTGCGAGAGTCTCGG 872
Dd 864 CGGCTTCGGGCGCTTCAGCAGCTGCAACAGCAGAGCGGCTGAGCAGCGCTGCGAGAGTCTCGG 923
Qy 873 GGAGACAGAGCGGTTCTTCCAGCAGCTGCTAGGCGCGGCTGCTGACAGCTGCTCCCGCAGG 932
Dd 924 GGAGACAGAGCGGTTCTTCCAGCAGCTGCTAGGCGCGGCTGCTGACAGCTGCTCCCGCAGG 983
Qy 933 CGAGGTGAGGTTCACAAAGATGAAGGCGCTGTACTTGGCCCTTGAAACCAAGTGCAGGTGAG 992
Dd 984 CGAGGTGAGGTTCACAAAGATGAAGGCGCTGTACTTGGCCCTTGAAACCAAGTGCAGGTGAG 1043

PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX

PS Claim 80; SEQ ID NO 410; 1762pp; English.

XX The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
CC viral, bacterial or fungal infection. The present sequence represents a
CC mRNA reference sequence for a 50 mer oligonucleotide marker for diagnosis
CC and monitoring of allograft rejection and other disorders.

XX Sequence 2700 BP; 451 A; 905 G; 850 G; 494 T; 0 U; 0 Other;

Query Match 98.2%; Score 2592.2; DB 12; Length 2700;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2599; Conservative 0; Mismatches 9; Indels 3; Gaps 1;

| | | | |
|----|-----|--|-----|
| QY | 33 | GGCGCGCGCAGCACACCCCGGAGACATGGGCTCCATGTTCCGGAGGAGGAGTGGCCCT | 92 |
| DB | 84 | GGCGCGCGCAGCACACCCCGGAGACATGGGCTCCATGTTCCGGAGGAGGAGTGGCCCT | 143 |
| QY | 93 | GGTCAGGCTTTTCTGCGCACAGCGCTGCCTACACCTGCGTAGCTCGGCTCGGCGAGCT | 152 |
| DB | 144 | GGTCAGGCTTTTCTGCGCACAGCGCTGCCTACACCTGCGTAGCTCGGCTCGGCGAGCT | 203 |
| QY | 153 | GGGCTCTGTGGAGTTTCAGAGACCTCAACGCCTCGGTGAGCGCTTCCAGAGACGCTTTGT | 212 |
| DB | 204 | GGGCTCTGTGGAGTTTCAGAGACCTCAACGCCTCGGTGAGCGCTTCCAGAGACGCTTTGT | 263 |
| QY | 213 | GTTTATGTTTGGCGCTGTGAGGAGCTGGAGAAGACCTTCACTTCTCTGCGAGGAGGT | 272 |
| DB | 264 | GTTTATGTTTGGCGCTGTGAGGAGCTGGAGAAGACCTTCACTTCTCTGCGAGGAGGT | 323 |
| QY | 273 | GGGCGGGCTGGGCTGTCTCCGCCCGCAAGGGAGGCTGCGGACACCCACCCCG | 332 |
| DB | 324 | GGGCGGGCTGGGCTGTCTCCGCCCGCAAGGGAGGCTGCGGACACCCACCCCG | 383 |
| QY | 333 | GGACTGTGCGCATCCAGGAGGAGACGGAGCGCTTGGCCAGGAGCTGGGGATGTGG | 392 |
| DB | 384 | GGACTGTGCGCATCCAGGAGGAGACGGAGCGCTTGGCCAGGAGCTGGGGATGTGG | 443 |
| QY | 393 | GGGCAACAGCAGGCGCTGCGGGCCAGCTGCACAGCTGACGCTCCAGCGCGGTGCT | 452 |
| DB | 444 | GGGCAACAGCAGGCGCTGCGGGCCAGCTGCACAGCTGACGCTCCAGCGCGGTGCT | 503 |
| QY | 453 | ACGCCAGGCGCATGAACCTCAGCTGGCAGCGCCACACAGATGGGCTCAGAGAGGAC | 512 |
| DB | 504 | ACGCCAGGCGCATGAACCTCAGCTGGCAGCGCCACACAGATGGGCTCAGAGAGGAC | 563 |
| QY | 513 | GCCCTGTCTCAGGCGCCCGGGGCGCGCACAGAACCTTGAAGGTCAACTTTGTGGCAGG | 572 |
| DB | 564 | GCCCTGTCTCAGGCGCCCGGGGCGCGCACAGAACCTTGAAGGTCAACTTTGTGGCAGG | 623 |
| QY | 573 | TGCGGTGAGCGCCACAGGCGCTGCTAGCGCGCTGCTGAGGGCTGCGCGG | 632 |
| DB | 624 | TGCGGTGAGCGCCACAGGCGCTGCTAGCGCGCTGCTGAGGGCTGCGCGG | 683 |
| QY | 633 | CTTCTCATTTGCGAGCTTCAGGAGCTGGAGCAGCGCTGGAGCACCCCGTGACGGCGGA | 692 |
| DB | 684 | CTTCTCATTTGCGAGCTTCAGGAGCTGGAGCAGCGCTGGAGCACCCCGTGACGGCGGA | 743 |
| QY | 693 | GCCAGCAGTGGATGACCTTCTCATCTCTTACTGGGGTGAGCAGATCGGACAGAGAT | 752 |
| DB | | | |

| | | | |
|----|------|---|------|
| DB | 744 | GCCAGCCAGCTGGATGACCTTCTCTCATCTCTCTACTGCGGTGAGCAGATCGGACAGAGAT | 803 |
| QY | 753 | CCGCAAGATCACGGAATGCTTCCACTGTCACACGCTCTTCCCGTTTCTGACAGAGAGGCG | 812 |
| DB | 804 | CCGCAAGATCACGGAATGCTTCCACTGTCACACGCTCTTCCCGTTTCTGACAGAGAGGCG | 863 |
| QY | 813 | CCGCTCTGGGGCCCTGACAGCTGCAACAGCAGAGACCGAGAGCTGCGAGAGTCTCTCGG | 872 |
| DB | 864 | CCGCTCTGGGGCCCTGACAGCTGCAACAGCAGAGACCGAGAGCTGCGAGAGTCTCTCGG | 923 |
| QY | 873 | GGAGACAGAGCGGTTCTTCTGAGCCAGTGCTAGCGCGGTGCTGAGCTGCTGCGCGCAGG | 932 |
| DB | 924 | GGAGACAGAGCGGTTCTTCTGAGCCAGTGCTAGCGCGGTGCTGAGCTGCTGCGCGCAGG | 983 |
| QY | 933 | GCAGTGCAGGTTCCAAAGATGAAGCCGTGTACTTGGCCCTGAAACCAAGTGCAGCGTGAG | 992 |
| DB | 984 | GCAGTGCAGGTTCCAAAGATGAAGCCGTGTACTTGGCCCTGAAACCAAGTGCAGCGTGAG | 1043 |
| QY | 993 | CACACGCGCACAGTGCGCTCATTTGCCAGGCGCTGCTGTCGCGAGGCTGCGCGCCCT | 1052 |
| DB | 1044 | CACACGCGCACAGTGCGCTCATTTGCCAGGCGCTGCTGTCGCGAGGCTGCGCGCCCT | 1103 |
| QY | 1053 | GCAGAGGCGCTTGGCGGACAGCTCGATGAGGAGGAGTGAGTGCCTGCTCACCGCAT | 1112 |
| DB | 1104 | GCAGAGGCGCTTGGCGGACAGCTCGATGAGGAGGAGTGAGTGCCTGCTCACCGCAT | 1163 |
| QY | 1113 | CCCTTGGCGGACATGCCCGCCACACTCATTCGCGACCAACCGCTTCAACCGGCTTCCA | 1172 |
| DB | 1164 | CCCTTGGCGGACATGCCCGCCACACTCATTCGCGACCAACCGCTTCAACCGGCTTCCA | 1223 |
| QY | 1173 | GGGCAATCGTGGATCGTACGGCGTGGCGGCTACAGGAGGTCACCCGCTCCCTACAC | 1232 |
| DB | 1224 | GGGCAATCGTGGATCGTACGGCGTGGCGGCTACAGGAGGTCACCCGCTCCCTACAC | 1283 |
| QY | 1233 | CATCATACCTTCCCTTCTCTGCTGTGATGTTTCGGGAGTGTGGGCGCTGCT | 1292 |
| DB | 1284 | CATCATACCTTCCCTTCTCTGCTGTGATGTTTCGGGAGTGTGGGCGCTGCT | 1343 |
| QY | 1293 | CATGTTCTTCTTGGCGCTGGCGCATGGTCTTTCGCGAGAACCGACCGGCTGTGAAAGCGCG | 1352 |
| DB | 1344 | CATGTTCTTCTTGGCGCTGGCGCATGGTCTTTCGCGAGAACCGACCGGCTGTGAAAGCGCG | 1403 |
| QY | 1353 | GCAGAAACAGAGATCTGGCAGACTTCTTTCAGGGGCGCTACCTGCTCTTATGGGCT | 1412 |
| DB | 1404 | GCAGAAACAGAGATCTGGCAGACTTCTTTCAGGGGCGCTACCTGCTCTTATGGGCT | 1463 |
| QY | 1413 | GTTCTCCATCTACACCGGCTTCTCATCAACAGAGTGTTCAGTTCGCGGACACAGCATCTT | 1472 |
| DB | 1464 | GTTCTCCATCTACACCGGCTTCTCATCAACAGAGTGTTCAGTTCGCGGACACAGCATCTT | 1523 |
| QY | 1473 | CCCTTGGGCTGGAGTGTGGCGCATGGCCCAACAGTCTGGCTGGAGTGATGCAATCT | 1532 |
| DB | 1524 | CCCTTGGGCTGGAGTGTGGCGCATGGCCCAACAGTCTGGCTGGAGTGATGCAATCT | 1583 |
| QY | 1533 | GGCCAGCACAGATGCTTACCTCGATCCCAACAGTCAACCGGTGCTTCTTGGGACCTTA | 1592 |
| DB | 1584 | GGCCAGCACAGATGCTTACCTCGATCCCAACAGTCAACCGGTGCTTCTTGGGACCTTA | 1643 |
| QY | 1593 | CCCTTTTGGCATCGATCTTATTTGGAGCTTGCTGTCGCAACCACTTGGAGTCTCTCAATC | 1652 |
| DB | 1644 | CCCTTTTGGCATCGATCTTATTTGGAGCTTGCTGTCGCAACCACTTGGAGTCTCTCAATC | 1703 |
| QY | 1653 | CTTCAAGATGAGATGCTCGCTCATCTGGGCGTGTGCAATGGGCTTTGGGCTGCT | 1712 |
| DB | 1704 | CTTCAAGATGAGATGCTCGCTCATCTGGGCGTGTGCAATGGGCTTTGGGCTGCT | 1763 |
| QY | 1713 | CGGAGCTTTCAACACAGTGCATTTTGGCAGAGGACCGGCTGCTGCTGGAGACGCTGCC | 1772 |
| DB | 1764 | CGGAGCTTTCAACACAGTGCATTTTGGCAGAGGACCGGCTGCTGCTGGAGACGCTGCC | 1823 |
| QY | 1773 | GGAGCTCACCTTCTGCTGGAGCTTTCGGTTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT | 1832 |
| DB | 1824 | GGAGCTCACCTTCTGCTGGAGCTTTCGGTTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT | 1883 |

333 GGACCTGTGCGCATCCAGAGGAGACGGAGGCGCTTGGCCCAAGGAGCTCGGGGATGTGG 392
Db GGACCTGTGCGCATCCAGAGGAGACGGAGGCGCTTGGCCCAAGGAGCTCGGGATGTGG 443
393 GGGCAACCAAGAGGCGCTCGGGGCCAGCTGCAACCAAGCTGCAAGCTTCAAGCGCGCTGT 452
Db GGGCAACCAAGAGGCGCTCGGGGCCAGCTGCAACCAAGCTGCAAGCTTCAAGCGCGCTGT 503
453 AGCCAGGGCCATGAACCTCAGCTGGCAGCGGCCCAACAGATGGGGCTCAGAGAGGAC 512
Db AGCCAGGGCCATGAACCTCAGCTGGCAGCGGCCCAACAGATGGGGCTCAGAGAGGAC 563
513 GCGCCCTGTCCAGGCGCCCGGGGGCGCACCAAGGACCTGAGGCTCAAATTGTGSCAGG 572
Db GCGCCCTGTCCAGGCGCCCGGGGGCGCACCAAGGACCTGAGGCTCAAATTGTGSCAGG 623
573 TGCCTGTGAGGCCCAAGAGGCCCTTGCCTAGAGCGCTGTCTTGGAGGGCTTGC CGCG 632
Db TGCCTGTGAGGCCCAAGAGGCCCTTGCCTAGAGCGCTGTCTTGGAGGGCTTGC CGCG 683
633 CTTCTCATTTGCGAGCTCAGGGAGCTGAGCAGCGCGCTGGAGCACCCCGTGAAGGGCA 692
Db CTTCTCATTTGCGAGCTCAGGGAGCTGAGCAGCGCGCTGGAGCACCCCGTGAAGGGCA 743
693 GCCAGCCAGTGGATGACCTTCTCATCTCTACTGGGGTGAGCAGATCGGACAGAAGAT 752
Db GCCAGCCAGTGGATGACCTTCTCATCTCTACTGGGGTGAGCAGATCGGACAGAAGAT 803
753 CGCGAAGATCAAGGACTGTTCATGCACTGCACTTCCGTTTCCGTTTCTGCAAGAGGAGGC 812
Db CGCGAAGATCAAGGACTGTTCATGCACTGCACTTCCGTTTCCGTTTCTGCAAGAGGAGGC 863
813 CGCGCTCGGGGCCCTGACAGCTGCAACAGCAGAGCCAGAGCTGCAAGAGGCTCTCGG 872
Db CGCGCTCGGGGCCCTGACAGCTGCAACAGCAGAGCCAGAGCTGCAAGAGGCTCTCGG 923
873 GGAGACAGAGCGGTTCTGAGCAGAGTGTAGCGCGGGTGTGCACTGCTGCGCGCAGG 932
Db GGAGACAGAGCGGTTCTGAGCAGAGTGTAGCGCGGGTGTGCACTGCTGCGCGCAGG 983
933 GAGGTGCAAGTTCACAGATGAAGGCGGTGTACTTGGCCCTGAAACCAAGTGCAGGTGAG 992
Db GAGGTGCAAGTTCACAGATGAAGGCGGTGTACTTGGCCCTGAAACCAAGTGCAGGTGAG 1043
993 CACCAAGCAAGTGCCTATTTGCGAGGCGCTGCTGTGCGAGACCTGCGCGCCCT 1052
Db CACCAAGCAAGTGCCTATTTGCGAGGCGCTGCTGTGCGAGACCTGCGCGCCCT 1103
1053 GCAGGAGGCCCTGCGGACAGCTCGATGGAGGAGGAGTGAAGTGGCTGCGGCTCAGCGCAT 1112
Db GCAGGAGGCCCTGCGGACAGCTCGATGGAGGAGGAGTGAAGTGGCTGCGGCTCAGCGCAT 1163
1113 CCCCTGCGGAGCATGCCCGCCCACTCATATCCGACCAACCGCTTCAAGGCGAGTTCA 1172
Db CCCCTGCGGAGCATGCCCGCCCACTCATATCCGACCAACCGCTTCAAGGCGAGTTCA 1223
1173 GGGCATCGTGATCGCTAGCGGCGTGGCGCTTACAGGAGGTCAACCCGCTCCCTACAC 1232
Db GGGCATCGTGATCGCTAGCGGCGTGGCGCTTACAGGAGGTCAACCCGCTCCCTACAC 1283
1233 CATCATCACTTCCCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1292
Db CATCATCACTTCCCTTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1343
1293 CATGTTCCTTTGCGCCCTGGCCATGCTTCTGCGAGAACCAACCGGCTGTGAAGCGCG 1352
Db CATGTTCCTTTGCGCCCTGGCCATGCTTCTGCGAGAACCAACCGGCTGTGAAGCGCG 1403
1353 GCAGAACAGAGATCTGGCAGCTTCTTCAAGGGGCGCTTACCTGCTTCTGCTTATGGGCT 1412
Db GCAGAACAGAGATCTGGCAGCTTCTTCTTCAAGGGGCGCTTACCTGCTTCTGCTTATGGGCT 1463

1413 GTTCTCCATCTACACCGGCTTCATCTACAAACGAGTGTCTTCACTCGCGCCACAGCATCTT 1472
Db GTTCTCCATCTACACCGGCTTCATCTACAAACGAGTGTCTTCACTCGCGCCACAGCATCTT 1523
1473 CCCCTCGGCTGGAGTGTGGCGCATGGCCAAACAGTCTTGGCTGGAGTGATGCAATTCCT 1532
Db CCCCTCGGCTGGAGTGTGGCGCATGGCCAAACAGTCTTGGCTGGAGTGATGCAATTCCT 1583
1533 GGCCAGCACACAGATGCTTACCTGGATCCCAACGTCAACCGGTGTCTTCTGGGACCCCTA 1592
Db GGCCAGCACACAGATGCTTACCTGGATCCCAACGTCAACCGGTGTCTTCTGGGACCCCTA 1643
1593 CCCCTTTGGCATCGATCTTATTGAGCTGTGCTGCAACCACTTGAAGTCTTCTCAACTC 1652
Db CCCCTTTGGCATCGATCTTATTGAGCTGTGCTGCAACCACTTGAAGTCTTCTCAACTC 1703
1653 TTCTAAGATGAAGATGTCCGTCACTCTGGGCGTCTGTGCAATGGGCTTTGGGTTGGTCT 1712
Db TTCTAAGATGAAGATGTCCGTCACTCTGGGCGTCTGTGCAATGGGCTTTGGGTTGGTCT 1763
1713 CGGAGTCTTCAACACAGTGCATTTGGCCAGAGGACACGGGCTGTCTGGAGACGCTGCC 1772
Db CGGAGTCTTCAACACAGTGCATTTGGCCAGAGGACACGGGCTGTCTGGAGACGCTGCC 1823
1773 GGAGCTCACTTCTGCTGGGACTCTTGGGTTACCTGCTGTCTTAGTCACTTCAAGTG 1832
Db GGAGCTCACTTCTGCTGGGACTCTTGGGTTACCTGCTGTCTTAGTCACTTCAAGTG 1883
1833 GCTGTGTGTGGGTTGCCAGGGCGGCTCG --- CCCAGCATCTTCACTTCACTCAA 1889
Db GCTGTGTGTGGGTTGCCAGGGCGGCTCG --- CCCAGCATCTTCACTTCACTCAA 1943
1889 CATGTTCTTCTTCTCCACAGCCCCCAGCAACAGGCTGTCTTACCCCGGAGAGGTGT 1949
Db CATGTTCTTCTTCTCCACAGCCCCCAGCAACAGGCTGTCTTACCCCGGAGAGGTGT 2003
1949 CCAGGCCAGCTGTGGTCTTGGGCTTGGCCATGTGTGCCATCTTGTCTTGGGACACC 2009
Db CCAGGCCAGCTGTGGTCTTGGGCTTGGCCATGTGTGCCATCTTGTCTTGGGACACC 2063
2063 CTTGCACTGTGTGCAACCGCCACCGCGCGCTGCGAGGAGGCGGCTGACCGACAGGA 2069
Db CTTGCACTGTGTGCAACCGCCACCGCGCGCTGCGAGGAGGCGGCTGACCGACAGGA 2123
2070 GGAAACAAAGCGCGGTTGCTGGACCTGCTGACGCACTGTGTAATGCTGGAGTCCGA 2129
Db GGAAACAAAGCGCGGTTGCTGGACCTGCTGACGCACTGTGTAATGCTGGAGTCCGA 2183
2183 TGAGGAAAGGCGAGGGGCTTGGATGATGAAGAGGAGGCGGAGCTCGTCCCTCGAGGT 2189
Db TGAGGAAAGGCGAGGGGCTTGGATGATGAAGAGGAGGCGGAGCTCGTCCCTCGAGGT 2243
2190 GCTCATGCAACAGGCGCATCCACCATCGAGTTCTGCTTGGGCTGCTTCCAAACCGCG 2249
Db GCTCATGCAACAGGCGCATCCACCATCGAGTTCTGCTTGGGCTGCTTCCAAACCGCG 2303
2250 CTTCTACCTGCGCTGTGGGCGCTGAGCCTGGCCACCGCCAGCTGCTCGAGGTTCTGTG 2309
Db CTTCTACCTGCGCTGTGGGCGCTGAGCCTGGCCACCGCCAGCTGCTCGAGGTTCTGTG 2363
2363 GGCCATGCTGATGCGCATAGGCTTGGGCGCGGAGGTGGGCTGCGGCTGTGTGT 2369
Db GGCCATGCTGATGCGCATAGGCTTGGGCGCGGAGGTGGGCTGCGGCTGTGTGT 2423
2370 GCTGTGCCCCATCTTTTGTGGCGCTTTTGGCGTATGACCGTGTGTATCTCTGCTGGTATGA 2429
Db GCTGTGCCCCATCTTTTGTGGCGCTTTTGGCGTATGACCGTGTGTATCTCTGCTGGTATGA 2483
2430 GGGAGCTCTAGCTTCTTCCGCGCTGCGGCTGCACTGGGTGGAAATCCAGAAACAGTT 2489
Db GGGAGCTCTAGCTTCTTCCGCGCTGCGGCTGCACTGGGTGGAAATCCAGAAACAGTT 2543
2490 CTACTCAGGCAAGGCTTACAAAGCTGAGTCCCTTCACTTCTGCTGCCACAGATGACTAGG 2549

| | | | |
|----|------|---|------|
| Db | 2544 | CTACTCAGGCACGGGCTACAAAGCTGAGTCCCTTCACCTTCGCTGCCACAGATGACTAGGG | 2603 |
| Qy | 2550 | CCCACTGCAGGTCTGCGACCTCCTTCTGACCTCTGAGGCAGGAGGGAATAAAGAC | 2609 |
| Db | 2604 | CCCACTGCAGGTCTGCGACCTCCTTCTGACCTCTGAGGCAGGAGGGAATAAAGAC | 2663 |
| Qy | 2610 | GGTCCGCCCTGGCAAAAAAAAAAAAAA | 2640 |
| Db | 2664 | GGTCCGCCCTGGCAGTCAAAAAAAAAAAAAA | 2694 |

Search completed: June 30, 2006, 04:27:43
Job time : 1513 secs

THIS PAGE BLANK (USPTO)

| | | | | | | | |
|----------------------------|------|--|------|--|------|--|------|
| Query Match | | | | 90.0%; Score 2376.6; DB 14; Length 2493; | | | |
| Best Local Similarity | | | | 95.9%; Pred. No. 0; | | | |
| Matches 2392; Conservative | | | | 0; Mismatches 98; Indels 3; Gaps 1; | | | |
| QY | 58 | ATGGGCTCATGTTCCGGAGCGAGGAGTGGCCCTGGTCCAGTCTTTCTGCCCCACAGCG | 117 | DB | 1021 | ATGGAGGAGGAGTGAAGTCCGTGCTACCGCATCCCTCGCGGACATGCCCCCACA | 1080 |
| DB | 1 | ATGGGCTCATGTTCCGGAGCGAGGAGTGGCCCTGGTCCAGTCTTTCTGCCCCACAGCG | 60 | QY | 1138 | CTCATCCGACCAACCGCTTTCACGCCAGCTTTCAGAGGGAATCGTGGATCGCTACCGCGTG | 1197 |
| QY | 118 | GCTGCTTACACCTGGTGAAGTGGCTGGGCGAGCTGGGCTCGTGGAGTTTCAGAGCCTC | 177 | DB | 1081 | CTCATCCGACCAACCGCTTTCACGCCAGCTTTCAGAGGGAATCGTGGATCGCTACCGCGTG | 1140 |
| DB | 61 | GCTGCTTACACCTGGTGAAGTGGCTGGGCGAGCTGGGCTCGTGGAGTTTCAGAGCCTC | 120 | QY | 1198 | GCGCGTACCAGGAGGTCAACCCGCTCCCTACACATCATCCTTCCCTTCCCTTCTT | 1257 |
| QY | 178 | AACGCTCCGCTGAGCGCTTCAGAGACGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 237 | DB | 1141 | GCGCGTACCAGGAGGTCAACCCGCTCCCTACACATCATCCTTCCCTTCCCTTCTT | 1200 |
| DB | 121 | AACGCTCCGCTGAGCGCTTCAGAGACGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT | 180 | QY | 1258 | GCTGTGATGTTTCGGGATGTGGGCCACCGGCTGCTCATGTTCTCTTCCCTCGGCCATG | 1317 |
| QY | 238 | CTGGAGAAACCTTCACTTCTGACAGGAGGTGCGCGGCTGGGCTGGTCTGCTGCCCC | 297 | DB | 1201 | GCTGTGATGTTTCGGGATGTGGGCCACCGGCTGCTCATGTTCTCTTCCCTCGGCCATG | 1260 |
| DB | 181 | CTGGAGAAACCTTCACTTCTGACAGGAGGTGCGCGGCTGGGCTGGTCTGCTGCCCC | 240 | QY | 1318 | GTCTTTCGGAGAACCGACCGGCTGTGAAAGCGCGCAGAACGAGATCTGGCAGACTTTC | 1377 |
| QY | 298 | CGGCCAAGGGAGGCTGCGGCAACCCCAACCCCGGACCTGCTGGCATCCAGAGGAG | 357 | DB | 1261 | GTCTTTCGGAGAACCGACCGGCTGTGAAAGCGCGCAGAACGAGATCTGGCAGACTTTC | 1320 |
| DB | 241 | CGGCCAAGGGAGGCTGCGGCAACCCCAACCCCGGACCTGCTGGCATCCAGAGGAG | 300 | QY | 1378 | TTACAGGGCGGCTACCTGCTCTGCTTATGGGCTGTCTCCATCTACACCGGCTTCATC | 1437 |
| QY | 358 | ACGGAGCGCTTGGCCAGAGCTGCGGATGTGCGGGCAACAGCAGGCGCTTGGCGGC | 417 | DB | 1321 | TTACAGGGCGGCTACCTGCTCTGCTTATGGGCTGTCTCCATCTACACCGGCTTCATC | 1380 |
| DB | 301 | ACGGAGCGCTTGGCCAGAGCTGCGGATGTGCGGGCAACAGCAGGCGCTTGGCGGC | 360 | QY | 1438 | TACAAACGAGTGTTCAGTCCGCCACACAGCATCTTCCCTCGGCTGGAGTGTGGCGGC | 1497 |
| QY | 418 | CAGCTGACACAGCTGAGCTCCACGCGCGTGTACGCCAGGGCCATGAACCTCAGCTG | 477 | DB | 1381 | TACAAACGAGTGTTCAGTCCGCCACACAGCATCTTCCCTCGGCTGGAGTGTGGCGGC | 1440 |
| DB | 361 | CAGCTGACACAGCTGAGCTCCACGCGCGTGTACGCCAGGGCCATGAACCTCAGCTG | 420 | QY | 1498 | ATGGCCAAACGAGTGTGGCTGAGTGAATGCAATTCCTGGGCCAGCACACGATGCTTACCCTG | 1557 |
| QY | 478 | GCAGCGGCCCAACAGATGGGCGCTCAGAGAGAGCGCCCTGCTCCAGGCCCGCGGGG | 537 | DB | 1441 | ATGGCCAAACGAGTGTGGCTGAGTGAATGCAATTCCTGGGCCAGCACACGATGCTTACCCTG | 1500 |
| DB | 421 | GCAGCGGCCCAACAGATGGGCGCTCAGAGAGAGCGCCCTGCTCCAGGCCCGCGGGG | 480 | QY | 1558 | GATCCCAACGTCACCGGTGTCTTCTGGGACCTTACCCTTGGGATCGATCGATCTTATGG | 1617 |
| QY | 538 | CGCCACACGAGCTCAGGGTCAACTTTGTGGAGGTGCGGTGGAGCGCCCAACAGGCGCT | 597 | DB | 1501 | GATCCCAACGTCACCGGTGTCTTCTGGGACCTTACCCTTGGGATCGATCGATCTTATGG | 1560 |
| DB | 481 | CGCCACACGAGCTCAGGGTCAACTTTGTGGAGGTGCGGTGGAGCGCCCAACAGGCGCT | 540 | QY | 1618 | AGCTGGCTGCCCAACCACTTGAAGTCTCTCAACTCTCTTCAAGATGAAGATGTCGCTCATC | 1677 |
| QY | 598 | GCCCTAGAGCGCTGCTGAGAGGCGTGGCCGCGCTTCTCATTTGCCAGCTTTCAGGGAG | 657 | DB | 1561 | AGCTGGCTGCCCAACCACTTGAAGTCTCTCAACTCTCTTCAAGATGAAGATGTCGCTCATC | 1620 |
| DB | 541 | GCCCTAGAGCGCTGCTGAGAGGCGTGGCCGCGCTTCTCATTTGCCAGCTTTCAGGGAG | 600 | QY | 1678 | CTGGCGCTGTGCAATGGCCCTTGGGGTGTGCTCGAGTCTTCAACGATGCACTTTT | 1737 |
| QY | 658 | CTGGAGCAGCGCTGAGACACCCGTGACGGGCGAGCGCACCGTGGATGACCTTCTC | 717 | DB | 1621 | CTGGCGCTGTGCAATGGCCCTTGGGGTGTGCTCGAGTCTTCAACGATGCACTTTT | 1680 |
| DB | 601 | CTGGAGCAGCGCTGAGACACCCGTGACGGGCGAGCGCACCGTGGATGACCTTCTC | 660 | QY | 1738 | GSCCAGAGGCAACCGCTGCTGAGAGCGCTGCGGAGCTCACCTTCTGCTGGGACTC | 1797 |
| QY | 718 | ATCTCTTACTGGGGTGAAGCAGATCGGACAGAGATCCGCAAGATCACGGACTGCTTCCAC | 777 | DB | 1681 | GSCCAGAGGCAACCGCTGCTGAGAGCGCTGCGGAGCTCACCTTCTGCTGGGACTC | 1740 |
| DB | 661 | ATCTCTTACTGGGGTGAAGCAGATCGGACAGAGATCCGCAAGATCACGGACTGCTTCCAC | 720 | QY | 1798 | TTCCGTTTACCTCGTGTCTTCTAGTCAATCTAAGTGGCTGTGTGTCTGGGCTGCCAGGGCC | 1857 |
| QY | 778 | TGCCAGCTTCCCGTTCTGACAGAGGAGGCGCGCTCGGGGCCCTGACGAGCTG | 837 | DB | 1741 | TTCCGTTTACCTCGTGTCTTCTAGTCAATCTAAGTGGCTGTGTGTCTGGGCTGCCAGGGCC | 1800 |
| DB | 721 | NN | 780 | QY | 1858 | GCCTG---CCGAGATCTCATCACTTCAATCAACATGTTCTCTTCTCCACAGCCCC | 1914 |
| QY | 838 | CAACAGCAGAGCGAGCTGAGGAGTCTCGGGAGACAGAGCGGTTCTGTAGCCAG | 897 | DB | 1801 | GCCTG---CCGAGATCTCATCACTTCAATCAACATGTTCTCTTCTCCACAGCCCC | 1860 |
| DB | 781 | NN | 840 | QY | 1915 | AGCAACGCTGCTTACCCCGGAGGAGTGTGTCAGGCGACCGCTGCTGGTCTCTGGCC | 1974 |
| QY | 898 | GTGCTAGGCGGGTGTGAGCTGCTGCGCGCAGGCGAGGTGAGTCCACAAGATGAAG | 957 | DB | 1861 | AGCAACGCTGCTTACCCCGGAGGAGTGTGTCAGGCGACCGCTGCTGGTCTCTGGCC | 1920 |
| DB | 841 | GTGCTAGGCGGGTGTGAGCTGCTGCGCGCAGGCGAGGTGAGTCCACAAGATGAAG | 900 | QY | 1975 | TTGGCCATGTGTGCCCATCTGCTGTGGCACACCCCTGCACTGCTGCACCGCACCCGC | 2034 |
| QY | 958 | GCCGTGTACTGCGCCCTGAACCAAGTGCAGCTGAGCAACGCAAAAGTGCCTCATGCC | 1017 | DB | 1921 | TTGGCCATGTGTGCCCATCTGCTGTGGCACACCCCTGCACTGCTGCACCGCACCCGC | 1980 |
| DB | 901 | GCCGTGTACTGCGCCCTGAACCAAGTGCAGCTGAGCAACGCAAAAGTGCCTCATGCC | 960 | QY | 2035 | GCGCCCTGCGGAGGAGCGCCGTGACCGACAGGAGGAAAACAGGCGGGTGTGTGGAC | 2094 |
| QY | 1018 | GAGGCTGTGTCTGTGCGAGACTGCGCGCCCTGCGAGGCGCTTGGGACAGCTCG | 1077 | DB | 1981 | GCGCCCTGCGGAGGAGCGCCGTGACCGACAGGAGGAAAACAGGCGGGTGTGTGGAC | 2040 |
| DB | 961 | GAGGCTGTGTCTGTGCGAGACTGCGCGCCCTGCGAGGCGCTTGGGACAGCTCG | 1020 | QY | 2095 | CTGCTGACGATCTGTAATGGCTGAGCTCCGATGAGGAAAAGGCGAGGGGCTTGGAT | 2154 |
| QY | 1078 | ATGGAGGAGGAGTGAAGTCCGTGCTCACCGCATCCCTGCGGGACATGCCCCCACA | 1137 | DB | 2041 | CTGCTGACGATCTGTAATGGCTGAGCTCCGATGAGGAAAAGGCGAGGGGCTTGGAT | 2100 |
| DB | | | | QY | 2155 | GATGAAGAGGAGCGGAGCTCGTCCCTCCGAGTGTCTCATGCAACGAGGCCATCCACACC | 2214 |

| | | | |
|----------------------------|--|--|---------------------|
| Db | 2101 | GATGAAGAGGAGCGGAGCTCGTCCCTCGAGGTGCTCATGCAACGAGGCATCCACACC | 2166 |
| Qy | 2215 | ATCAGATTGCGCTGGGCTGGCTCTCAACACACCGCCCTCTACCTGCGCCTGTGGGCGCCTG | 2274 |
| Db | 2161 | ATCAGATTGCGCTGGGCTGGCTCTCAACACACCGCCCTCTACCTGCGCCTGTGGGCGCCTG | 2220 |
| Qy | 2275 | AGCCTGGCCACGCGCCAGCTGTCGAGAGTTCTGTGGGCGCATGCTGATGCGCATAGGCGCTG | 2334 |
| Db | 2221 | AGCCTGGCCACGCGCCAGCTGTCGAGAGTTCTGTGGGCGCATGCTGATGCGCATAGGCGCTG | 2280 |
| Qy | 2335 | GGCCTGGCGCGGAGGTGGGCGTGGCGGCTGTGGTGTGGTCCCATCTTTGGCGCCTTT | 2394 |
| Db | 2281 | GGCCTGGCGCGGAGGTGGGCGTGGCGGCTGTGGTGTGGTCCCATCTTTGGCGCCTTT | 2340 |
| Qy | 2395 | GCCGTGATGACCGTGGCTATCTCTGCTGGTGATGAGGAGCTCTCAGCCTTCCCTGCAGGCC | 2454 |
| Db | 2341 | GCCGTGATGACCGTGGCTATCTCTGCTGGTGATGAGGAGCTCTCAGCCTTCCCTGCAGGCC | 2400 |
| Qy | 2455 | CTGGGCTGCACCTGGTGGAAATTCAGAAACAAGTTCTACTCAGCAGCGGCTCAAGCTG | 2514 |
| Db | 2401 | CTGGGCTGCACCTGGTGGAAATTCAGAAACAAGTTCTACTCAGCAGCGGCTCAAGCTG | 2460 |
| Qy | 2515 | AGTCCCTTACCTTCGCTGCGTGGCCACAGATGACTAG | 2547 |
| Db | 2461 | AGTCCCTTACCTTCGCTGCGTGGCCACAGATGACTAG | 2493 |
| RESULT 2 | | | |
| AY407800 | | | |
| LOCUS | AY407800 | 2493 bp DNA linear | GSS 15-DEC-2003 |
| DEFINITION | Pan troglodytes TCR1G1 gene, VIRTUAL TRANSCRIPT, partial sequence, Genomic survey sequence. | | |
| ACCESSION | AY407800 | | |
| VERSION | AY407800.1 | GI:39763771 | |
| KEYWORDS | GSS. | | |
| SOURCE | Pan troglodytes (chimpanzee) | | |
| ORGANISM | Pan troglodytes | | |
| REFERENCE | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios Science 302 (5652), 1960-1963 (2003) | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | |
| PUBMED | 14671302 | | |
| REFERENCE | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. Direct Submission Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence was made by sequencing genomic exons and ordering them based on alignment. | | |
| TITLE | Location/Qualifiers | | |
| JOURNAL | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | |
| COMMENT | This sequence was made by sequencing genomic exons and ordering them based on alignment. | | |
| FEATURES | | | |
| source | Location/Qualifiers | | |
| | 1. .2493 | | |
| | /organism="Pan troglodytes" | | |
| | /mol_type="genomic DNA" | | |
| | /db_xref="taxon:9598" | | |
| gene | <1_->2493 | | |
| | /gene="TCIRG1" | | |
| | /locus_tag="HCM3011" | | |
| ORIGIN | | | |
| Query Match | 76.2% | Score 2011.8; | DB 14; Length 2493; |
| Best Local Similarity | 81.5% | Pred. No. 0; | |
| Matches 2032; Conservative | 0; Mismatches 458; Indels 3; Gaps 1; | | |

```
1081 CTATCCGACCAACCGCTTACGCGCCAGCTTCCAGGGCATCGTGGATCCTACGCGGTG 1140
1198 GGCCTGTACAGAGGTCAACCGCGTCCCTTACACATCATCACTTCCCTTCTGTGTTT 1257
1141 GGCCTGTACAGAGGTCAACCGCGTCCCTTACACATCATCACTTCCCTTCTGTGTTT 1200
1258 GCTGTGATGTTGCGGATGTGGGCACCGGCTGTCTATGTTCTCTTCCGCTTGGCCATG 1317
1201 GCTGTGATGTTGCGGATGTGGGCACCGGCTGTCTATGTTCTCTTCCGCTTGGCCATG 1260
1318 GTCTTGGCGAGAACCGACCGGCTGTGAAGCGCGCAGAACGAGATCTGGCAGACTTTC 1377
1261 GTCTTGGCGAGAACCGACCGGCTGTGAAGCGCGCAGAACGAGATCTGGCAGACTTTC 1320
1378 TTACGGGGCGGTACTCTGTCTTCTGTATGGGCTGTCTTCTCATCATACACGGCTTCATC 1437
1321 TTACGGGGCGGTACTCTGTCTTCTGTATGGGCTGTCTTCTCATCATACACGGCTTCATC 1380
1438 TACAAGAGGTGTTCAAGTCGCGCCACACGATCTTCCCTCGGGCTGGAGTGTGGCGGC 1497
1381 TACAAGAGGTGTTCAAGTCGCGCCACACGATCTTCCCTCGGGCTGGAGTGTGGCGGC 1440
1498 ATGGCCAAACAGTCTGGCTGGAGTGAATCTCTGGCCACACACGATGCTTACCCCTG 1557
1441 ATGGCCAAACAGTCTGGCTGGAGTGAATCTCTGGCCACACACGATGCTTACCCCTG 1500
1558 GATCCCAAGCTCACCGGTGTCTCTGGGACCTACCCCTTGGGATCGATCTATTTGG 1617
1501 GATCCCAAGCTCACCGGTGTCTCTGGGACCTACCCCTTGGGATCGATCTATTTGG 1560
1618 AGCTGGCTGCCAACACATGAGCTTCTCAACTCTTCAAGTGAAGATGTCGCTCATC 1677
1561 AGCTGGCTGCCAACACATGAGCTTCTCAACTCTTCAAGTGAAGATGTCGCTCATC 1620
1678 CTGGGCGTGTGCACATGCGCTTTGGGGTGGTCTCTGGAGTCTTCAACACGATGCACTTT 1737
1621 CTGGGCGTGTGCACATGCGCTTTGGGGTGGTCTCTGGAGTCTTCAACACGATGCACTTT 1680
1738 GGCACAGGACCGGTGTCTGTGGAGCGTGTCCGGAGCTCACCTTCTGTCTGGGATC 1797
1681 GGCACAGGACCGGTGTCTGTGGAGCGTGTCCGGAGCTCACCTTCTGTCTGGGATC 1740
1798 TTGGTGTACTCTGTGTCTAGTCACTACAGTGGCTGTGTGTCTGGGCTGCCAGGGCC 1857
1741 TTGGTGTACTCTGTGTCTAGTCACTACAGTGGCTGTGTGTCTGGGCTGCCAGGGCC 1800
1858 GCTCTG---CCGACATCTCATCCATTCATCAATGTTCTCTTCTTCCACAGCCCC 1914
1801 GCTCTGCGCCCGCAGCATCTCATCCATTCATCAATGTTCTCTTCTTCCACAGCCCC 1860
1915 AGCAACAGGCTGCTTACCCCGGAGGAGGTGTCCAGGCCACGCTGTGTGTCTGGCC 1974
1861 AGCAACAGGCTGCTTACCCCGGAGGAGGTGTCCAGGCCACGCTGTGTGTCTGGCC 1920
1975 TTGGCCATGGTCCCATCTGTCTGTGTGGCACACCCCTGCACCTGTGCACCGCCACCGC 2034
1921 TTGGCCATGGTCCCATCTGTCTGTGTGGCACACCCCTGCACCTGTGCACCGCCACCGC 1980
2035 CGCCGCTGTGGAGAGGCGCGTGTGACCGACAGGAGGAAAAAAGGCGGGTGTGTGGAC 2094
1981 CGCCGCTGTGGAGAGGCGCGTGTGACCGACAGGAGGAAAAAAGGCGGGTGTGTGGAC 2040
2095 CTGCTGTAGCGATCTGTGAATGGCTGGAGTCCGATGAGGAAAAAGGCGGGCTGTGAT 2154
2041 CTGCTGTAGCGATCTGTGAATGGCTGGAGTCCGATGAGGAAAAAGGCGGGCTGTGAT 2100
2155 GATGAAGAGGAGGCGGAGCTCGTCTCCCTCCGAGGTGCTCATGCACAGGAGGATCCACACC 2214
2101 GATGAAGAGGAGGCGGAGCTCGTCTCCCTCCGAGGTGCTCATGCACAGGAGGATCCACACC 2160
2215 ATCGAGTTCTGCTGGGCTGTCTTCCAAACAACCGGCTCTTACCGGCTGTGTGGGCGCTG 2274
```

```
2161 ATTGAGTTCTGCTGCGCTGGCTCTTCCAAACACCGCTCTCTACCTTGGCCCTGTGGGCGCTG 2220
2275 AGCCTGGGCCACCGCCAGCTGTCCAGAGTTCTGTGGGCCATAGTGATGGCGATAGGCGCTG 2334
2221 AGCCTGGGCCACCGCCAGCTGTCCAGAGTTCTGTGGGCCATAGTGATGGCGATAGGCGCTG 2280
2335 GGCCTGGGCGCGGAGGTGTGGGCGTGGGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 2394
2281 GGCCTGGGCGCGGAGGTGTGGGCGTGGGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 2340
2395 GCGGTGATGACCGTGGCTATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2454
2341 GCGGTGATGACCGTGGCTATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2400
2455 CTGCGGCTGCACTGGGTGGATTCAGAAACAGATTTCTACTCAGGACAGGCGCTACAAGCTG 2514
2401 CTGCGGCTGCACTGGGTGGATTCAGAAACAGATTTCTACTCAGGACAGGCGCTACAAGCTG 2460
2515 AGTCCCTTCACTTCCGCTTCCGCTGCCACAGATGACTAG 2547
2461 AGTCCCTTCACTTCCGCTTCCGCTGCCACAGATGACTAG 2493
```

CR594613 1864 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DD004Y118 of Neuroblastoma Cot
50-normalized of Homo sapiens (human).

CR594613
HTC; CNSLT cDNA.
http://50475420
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE
1 (bases 1 to 1864)
Li W.B., Gruber C., Jessee J. and Polayes D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue
Genoscope.
2 (bases 1 to 1864)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES
Location/Qualifiers
1..1864
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD004Y118"
/tissue type="Neuroblastoma Cot 50-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 68.9%; Score 1818; DB 6; Length 1864;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1834; Conservative 0; Mismatches 5; Indels 3; Gaps 1;
QY 768 CTGCTTCCATGCCACGCTTCCCTTCTGTGACAGAGGAGGCGCGCTCGGCGCCCT 827
Db 23 CAGCTTCCATGCCACGCTTCCCTTCTGTGACAGAGGAGGCGCGCGCTCGGCGCCCT 82
QY 828 GCACAGCTGCACACAGACGACGAGCTGTGACGAGGTCTCTCGGGGAGACAGACGCGTT 887

| | | |
|---|----------------------|--|
| 3 | REFERENCE AUTHORS | Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Tashiro, H., Itoh, M., Sum, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Tashiro, H., Itoh, M., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. |
| | TITLE | RIKEN integrated sequence analysis (RISA) system-384-format |
| | JOURNAL | Genome Res. 10 (11), 1757-1771 (2000) |
| | PUBMED | 11076861 |
| 4 | REFERENCE AUTHORS | Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, D., Washio, T., Sakai, K., Okido, I., Furuno, M., Aono, H., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y. |
| | TITLE | RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium |
| | JOURNAL | Nature 409 (6821), 685-690 (2001) |
| | PUBMED | 11217851 |
| 5 | REFERENCE AUTHORS | Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Oato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chochis, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reid, J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Iehii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sasaki, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y. |
| | TITLE | Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs |
| | JOURNAL | Nature 420 (6915), 563-573 (2002) |
| | PUBMED | 12466851 |
| 6 | REFERENCE AUTHORS | Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzi, R., |
| | TITLE | Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome |
| | JOURNAL | Science 309 (5740), 1559-1563 (2005) |
| | PUBMED | 16141072 |
| | REFERENCE AUTHORS | Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Paghini, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C. |
| | TITLE | RIKEN Genome Exploration Research Group |
| | JOURNAL | Antisense transcription in the mammalian transcriptome |
| | PUBMED | 16141073 |
| | REFERENCE AUTHORS | 8 (bases 1 to 3031) |
| | TITLE | Direct Submission |
| | JOURNAL | Submitted (30-MAR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@sc.riken.jp, URL: http://genome.fgc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) |
| | COMMENT | cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. |

Trust/MRC building Addenbrookes Hospital (Cambridge) whose assistance we gratefully acknowledge.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>

FEATURES

source
1. 3031
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:F630203K02"
/db_xref="taxon:10090"
/clone="F630203K02"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
misc_feature
1. 3031
/note="T-cells, immune regulator 1 (MGD|MG1:1350931
GB|AB022322, evidence: BLASTN, 99%, match=2571)
putative"

ORIGIN

Query Match 65.5%; Score 1729; DB 6; Length 3031;
Best Local Similarity 82.1%; Pred. No. 0;
Matches 2080; Conservative 0; Mismatches 430; Indels 23; Gaps 7;
QY 40 GCAGCACACCGGGGACCATGGCTCCATGTTCCGGAGCGAGGAGTGGCCCTGGTCCAG 99
DB 448 GTAACAAATCAAGGATCATGGGCTCTATGTTCCGGAGTGAAGAGTGGGCTTGGTCCAG 507
QY 100 CTCTTTCTGCCACAGCGGCTGCTACACCTGGTGTAGTGGCTGGGCGAGCTGGGCTC 159
DB 508 CTCTCTGTTGCCACAGGCTCTGTTCAACTGTGTAGCCAGCTTGTGTAGCTGGGCTC 567
QY 160 GTGGAGTTACAGACCTCAACGCTCTGGTGAGCGCTTCCAGAGACCTTTGTGGTTGAT 219
DB 568 GTGGAGTTACAGACCTCAACGAATCGTGAGCGCTTCCAGAGACGCTTTCGTTGGTAGAC 627
QY 220 GTTGGGCTGTAGAGCTGGAGAGACCTTACCTTCTTCAGAGAGAGTGGCGCG 279
DB 628 GTCCGGCGTGCAGGAACTGGAGAGACGTTTACCTTCTTGGGGAGAAAGTGCACGG 687
QY 280 GCTGGGCTGCTGCCCGCCGCAAGGGGAGGCTGCGGCAACCCCGCCCGGACCTG 339
DB 688 GCAGGTCTGACGCTGGCCCGCCACCTGAGGGGACACTGCCAGCACCCCTCCCGAGACTG 747
QY 340 CTCGCGATCAGAGGAGACGAGCGCTTGGCCAGAGCTGCGGAGTGTGCGGGGCAAC 399
DB 748 CTCGCGATCAGAGGAAACGGATCGCTTGGCCAGAGCTTCCGGATGTTCTGGCAAT 807
QY 400 CAGAGGCTCTGGGGCCAGCTGACAGCTGCGGCTTCCAGAGCTTCCAGCGCGCTGCTACGCCAG 459
DB 808 CAGAGGCACTGGGGGCCAGCTGCAAGCTGCGGCTTCACTCGGCTGTGCTGGGCGAG 867
QY 460 GGCATGAACTCAGCTGGCAGCGCCGCAACAGATGGGCG ---CTCAGAGAGGAGCGCC 516
DB 868 AGCCACAGCCCGGCTGGCAGCTGACCACTATGAAGACCTTCTCTGAGACAAACACC 927
QY 517 CTGCTCCAGGCCCGCGGGGCGCCAGCAGGACCTGAGGCTCAACTTTGTGGCAGGTGCC 576
DB 928 CTGCTTCCAGGCACTCCGGGGGCCACATTCAGACCTGAAGTCAATTTGTGGCAGGTGCT 987
QY 577 GTGGAGCCCAAGGGCCCTGCGCTTAGAGCGCTGCTGTGGAGGGCTGCGCGGGCTTC 636
DB 988 GTGGAGCCCTTACAAGGCTGCGGCGCTGGAGCGCTGCTGTGGAGAGCTTCCCGGGCTTC 1047
QY 637 CTCATTGGCAGCTTCAGGAGCTGGAGAGCGCCCTGGAGCAGCCCGCTGAGCGGGAGCCCA 696
DB 1048 CTCATTGGCAGCTTTCGGGGAGACTGAGGGGAGCTGGAGGACCCAGTGCAGGCTGAGCCT 1107
QY 697 GCCAGCTGAGTACCTTCTCATCTCTCTAGCGGCTGAGCAGATCGGAGAGAGATCCGC 756
DB 1108 GCAACTTGGATGACCTTGTCTCTCTCTACTTGGGGTGAGCAGATTGGACAGAGATCCGA 1167

QY 757 AAGATCAGGACCTGCTTCCATCGCAGCTCTTCCCGTTTCTGCGAGGAGGAGGCGCCG 816
DB 1168 AAGATCAGGACCTGCTTCCATCGCAGCTCTTCCCGTTTCTGCGAGGAGGAGGCGCCG 1227
QY 817 CTGGGGCCCTGCGAGGAGCTGCAACAGCAGCAGCAGGAGCTGCAAGGAGTCTCTCGGGAG 876
DB 1228 TTCAGGACCTTGCAGCAGCTGCAGCAGCAAGCAGGAGCTTCAAGGAGTCTCTGGGAG 1287
QY 877 ACAGAGCGTTCCTGAGCCAGGTGCTAGGCCGGGTGCTGAGCTGTCTGCCGCCAGGGCAG 936
DB 1288 ACAGAGCGTTCCTGAGCCAGGTGTTGGGCGGGTGCAGCAGCTGTCTGCCCGCCATGCGAG 1347
QY 937 GTGAGGTTCACAAGATGAAGGCGGTGTACTTGGCCCTGAAACCAAGTGCAGGCTGAGCACC 996
DB 1348 GTGAGATTCACAAGATGAAGGCGGTGTACTTGAACCTCAACCAAGTGCAGTGTGAACACC 1407
QY 997 ACCCAGAGTGCCTCATTTGCCGAGGCTGTGTCTGTGCGAGACCTGCGCGCCCTGCGAG 1056
DB 1408 ACACAGGTGCTCTATCGCGGAGGTCTGTGTGCGCGAGGAGACCTGCGCCACTGTGCGAG 1467
QY 1057 GAGGCCCTTGGCGGACAGCTCGATGAGGAGGAGGTGAGTGCCTGGCTCACCGCATCCCC 1116
DB 1468 CAGCGCTGCAGAGCGCTCAAGTGAAGGAGGAGTGTGTCTGTCTCACCGCATCCCC 1527
QY 1117 TGGCGGAGCATGCCCGCCCACTATCCGACCAACCGCTTTCAGGCGCAGCTTCCAGGGC 1176
DB 1528 TGCAGGACATGCTCCAACTTCATCAGAGCAACCGCTTTCAGCTCCAGCTTCCAGGGC 1587
QY 1177 ATGCTGATCGCTACCGGCTGGGCGCTACACAGAGTCAACCGCTTCCCTACACCATC 1236
DB 1588 ATTGTGATGCCCTATGTTGTGGCGCTCAGGAGTTAACTCTGCTCCCTACACCATC 1647
QY 1237 ATCACTTCCCTTCTCTTGTGTGTGTTGTTGCGGAGTGTGGGCCACGCGCTGCTCATG 1296
DB 1648 ATCACTTCCCTTCTCTTGTGTGTGTTGTTGCGATGTTTGGGATGTTGGGCGACGACTGCTCATG 1707
QY 1297 TTCTCTTGGCCCTGGCCATGCTTTCAGGGCGCTTACCTGCTCTGCTTATGGGCTGCTTC 1356
DB 1708 TTTCTCTTGGCCCTGGCCATGCTTCTCACTGAACCGCTCAGCTGTGAAGGCTGCACAA 1767
QY 1357 AACGAGATCGGCGAGCTTCTTTCAGGGCGCTTACCTGCTCTGCTTATGGGCTGCTTC 1416
DB 1768 AACGAGATCGGCGAGCTTCTTTCAGGGCGCTGATCTCTCTGCTCATGGGCTGCTTC 1827
QY 1417 TCCATCTACACCGCTTCTATCAACAGTGTCTCAGTGCAGCGCCACAGCAGCTTCTCCCC 1476
DB 1828 TCGCTCTACACTGCTTCTATCACTAGTGTCTCAGCGCGAGCCACCACTATTTTCCCC 1887
QY 1477 TCGGCTGAGTGTGGCGCCATGGCCCAACAGCTGTGGCTGGAGTGAATGCTCTGGCC 1536
DB 1888 TCAGGTTGGAGTGTGGCTGCATGGCCCAACAGCTCAGGCTGGAGTGCAGGATATCTGTCC 1947
QY 1537 CAGCAGAGTGTACCTTACCTGGATCCCAAGTCAACCGGTGCTTCTTGGGAGCCCTA-CCC 1595
DB 1948 CAGCATCTCATGCTCACCCTGAACCTTAACATCACTGTGTGTCTTCTTGGACCATATCCC 2007
QY 1596 CTTTGGCATCGATCTTATTTGGAGCTGGCTGCAACCACTTGAAGTCTTCTCAACTCCTT 1655
DB 2008 TTTTGGCATTTGACCCGATCTGGAGCTTGGCCCAAGACCACTGAGCTTCTCAACTCCTT 2067
QY 1656 CAAGATGAAGTGTGCTGAT-CTTGGGCTGTGTGACATGAGCTTGGGGTGTGCTCTG 1714
DB 2068 CAAGATGAAGTGTGTCTATCTCTTGGGCTACGACATGAGCTTGGGGTGTGCTCTCA 2127
QY 1715 GAGTCTTCAACAGCTGCACTTTGGCCAGAGGACCGGCTGCTGCTGGAGACGCTGCCG 1774
DB 2128 GCATCTTCAACAGCTGCACTTTGGCCAGGCCACCGGCTACTGCTGGAGACCTTGGCTG 2187
QY 1775 AGCTCACTCTTCTGCTGGGACTTCTCGGTTTACCTCTGTTTCTTAGTCACTTACAAGTGGC 1834
DB 2188 AGCTCATCTCTGCTGGGCTCTCTTGGCTTACCTTGTGTTCTCATGTCTACAAGTGGG 2247
QY 1835 TGTGTGTGTGGGCTGCCAGGGCGGCTG ---CCAGCATCTCTCATCCACTTCAACA 1891

| | | | |
|----|------|--|------|
| Db | 2248 | TGAATGTCAGCTGCTAGGCGCTCTCTGGGCCCGGTAATCTCTATTCATTCTATCAATA | 2307 |
| Qy | 1892 | TGTTCTCTTCTCCACAGCCCAAGCAACAGGCTGCTTACCCGCCGACGAGAGTGCTCC | 1951 |
| Db | 2308 | TGTTCTCTTCTCTCAAAAACCCACCAATCATCTGCTCTTCCATGGGCGAGAGTGCTAC | 2367 |
| Qy | 1952 | AGGCCACGCTGGTGGTCTCTGGCTTTGGCCATATGFGGCCCATCTCTGCTCTTTGGCACACCCC | 2011 |
| Db | 2368 | AGTACGTGCTGGTGGTCTCTGGCTTTGGCTACCGTTCTCTATCTCTGCTGGGCACACCT | 2427 |
| Qy | 2012 | TGCACCTGCTGCACCGCCACCGCCGCG-----CCTGGGAGAGAGCCCGCTGACCGAC | 2065 |
| Db | 2428 | TGTACCTGCTGGCGCCAGCACCGCCACAGAAACAATCTAGAGAAGGCCACGAGCCGACG | 2487 |
| Qy | 2066 | AGGAGGAAAAACAAGCGCGGTGCTGTGACCTGCTGACGCATC-----TGTGAATGGCT | 2119 |
| Db | 2488 | AGGATGAGGACACGGACAAGCTTCTGGCTCCCTGTATGTCATCCACCTTGGAGAACAGCT | 2547 |
| Qy | 2120 | GGAGCTCCGATGAGGAAAAGCAGGGGGCTTGGATGATGAGAGGAGGCCGAGCTCGTCC | 2179 |
| Db | 2548 | GGAGCCCTGATGAGGAGAAGGCTGGGAGCCCGGGGATGAA---GAAACTGAGTTTGTCC | 2604 |
| Qy | 2180 | CCTCCGAGGTGCTCATGCACAGGCCATCCACACCATCGAGTTCTGCTCGGCTGCGCTCT | 2239 |
| Db | 2605 | CTTCTGAGATCTTCATGSCACAGCAATCCACACCATTGAGTTCTGCTGGGCTGCATCT | 2664 |
| Qy | 2240 | CCAAACCGGCTCTACTCTGGGCTGTGGGCTGTAGCCTTGGGCCACCGCCAGCTGTCCG | 2299 |
| Db | 2665 | CCAAACAGGCTCTACTTTCGCTCTCTGGGCGCTGAGCCTGGGCCCATGCCCAGCTGTCTG | 2724 |
| Qy | 2300 | AGGTTCTGTGGGCCATGGTATGGCATAGGCCTGGGCTGGGCGGGAGGTGGGCGTGG | 2359 |
| Db | 2725 | AGGTCTGTGGGCCATGGTATGGCATAGGCTTGGGCGTGGGCGGAGATGGGCGTGG | 2784 |
| Qy | 2360 | CGGCTGTGTGCTGCTCCCATCTTTTGGCGCTTTTGGCGTATGACCGTGGCTATCTCGC | 2419 |
| Db | 2785 | CAGCTGTGGTGTGCTCCCTGTGTTTGTGCTGCCCTTGTGTGTGATCTGGCCATCTGT | 2844 |
| Qy | 2420 | TGTTGATGGAGGAGCTCTCAGCCTTCTGACGCGCTTGGGCTGCACTGGGTGGAAATTC | 2479 |
| Db | 2845 | TAGTGATGGAGGGCTCTCAGCCTTCTGTCATGCCCCCTGGGCTTGGCACTGGGCTGGAGTTC | 2904 |
| Qy | 2480 | AGAAACAGTTCTACTCAGGACAGGGCTACAAGCTGAGTCTCCTTACCTTTCGCTGCCACAG | 2539 |
| Db | 2905 | AGAAACAGTTCTATTAGGACACCGGCTACAAGCTACAGCTCACCCCTTCACTGTGGACA | 2964 |
| Qy | 2540 | ATGACTAGGCCCC | 2552 |
| Db | 2965 | GTGACTAGCTACC | 2977 |

| | |
|------------|---|
| RESULT 5 | |
| AY407801 | |
| LOCUS | 2514 bp DNA linear GSS 15-DEC-2003 |
| DEFINITION | Mus musculus TcIRG1 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence. |
| ACCESSION | AY407801.1 GI:39763772 |
| VERSION | GSS. |
| KEYWORDS | Mus musculus (house mouse) |
| SOURCE | Mus musculus |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidae; Muridae; Murinae; Mus. 1 (bases 1 to 2514) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios |
| TITLE | Science 302 (5652), 1960-1963 (2003) |
| JOURNAL | |

COMMENT - Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES

source
Location/Qualifiers
1..1655
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1059YF06"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 55.6%; Score 1468.6; DB 6; Length 1655;
Best Local Similarity 94.9%; Pred. No. 3.1e-274;
Matches 1571; Conservative 0; Mismatches 4; Indels 80; Gaps 2;
QY 1022 CCTGTGCTCTGTGGAGACCTGCCGCCCTGCAGAGGCCCTGCGGACAGCTCGATGG 1081
DB 1 CCTGTGCTCTGTGGAGACCTGCCGCCCTGCAGAGGCCCTGCGGACAGCTCGATGG 60
QY 1082 AGGAGGAGTGTAGTGCCTGCTACCGCATCCCTGCGGAGCATGCCCCCACTCA 1141
DB 61 AGGAGGAGTGTAGTGCCTGCTACCGCATCCCTGCGGAGCATGCCCCCACTCA 120
QY 1142 TCGCACCAACCGCTTCAGGCGAGCTTCAGGGCATCGTGTAGTGTAGGCGG 1201
DB 121 TCGCACCAACCGCTTCAGGCGAGCTTCAGGGCATCGTGTAGTGTAGGCGG 180
QY 1202 GCTACAGGAGTCAACCCCGCTCCCTACACCATCATACCTTCCTTCCTGTTGCTG 1261
DB 181 GCTACAGGAGTCAACCCCGCTCCCTACACCATCATACCTTCCTTCCTGTTGCTG 240
QY 1262 TGATGTTGCGGATGTGGGCCACGGCTGCTCATGTTCTTCCTGCGCCCTGGCCATGGTCC 1321
DB 241 TGATGTTGCGGATGTGGGCCACGGCTGCTCATGTTCTTCCTGCGCCCTGGCCATGGTCC 300
QY 1322 TTGCGGAGAACCGAGCGGTGTGAAGCGCGCAGAACGAGATCTGGCAGCTTCTTCA 1381
DB 301 TTGCGGAGAACCGAGCGGTGTGAAGCGCGCAGAACGAGATCTGGCAGACTTCTTCA 360
QY 1382 GGGGCGCTTACCTGCTCTGCTTATGGGCTGTCTTCCATCTACACCGGCTTCATCTACA 1441
DB 361 GGGGCGCTTACCTGCTCTGCTTATGGGCTGTCTTCCATCTACACCGGCTTCATCTACA 420
QY 1442 ACGAGTGTTCAGTTCGCGCCACAGCATCTTCCCTGCGGCTGGAGTGTGGCGCATGG 1501
DB 421 ACGAGTGTTCAGTTCGCGCCACAGCATCTTCCCTGCGGCTGGAGTGTGGCGCATGG 480
QY 1502 CCAACGAGTCTGGCTGG----- 1518
DB 481 CCAACGAGTCTGGCTGGAGGTGAGGCCCGGGCCCCAGCGCGCTGGGGCCCCCGCAGCAC 540
QY 1519 -----AGTGATGATTCCTGGCGAGTGTGATTCCTGGCCAGCACAC 1544
DB 541 CGCAGCCCTGACCGCCCTCCCTGGGTGGCGGAGTGTGATTCCTGGCCAGCACAC 600
QY 1545 GATGTTACCTGGATGCCAAGTCAACGGTGTCTTCCTGGGACCTACCTCCCTTTGGCAT 1604
DB 601 GATGTTACCTGGATGCCAAGTCAACGGTGTCTTCCTGGGACCTACCTCCCTTTGGCAT 660
QY 1605 CGATCCTATTTGGAGCGCTGGCCGACACCATGTAGCTTCCTCAACTCTTCAAGTGA 1664
DB 661 CGATCCTATTTGGAGCGCTGGCCGACACCATGTAGCTTCCTCAACTCTTCAAGTGA 720
QY 1665 GATGTCCTGTCATCTGGGCGTGTGTCACATGGCCCTTTGGGCTGTCTCGAGTCTTCAA 1724
DB 721 GATGTCCTGTCATCTGGGCGTGTGTCACATGGCCCTTTGGGCTGTCTCGAGTCTTCAA 780
QY 1725 CCACGTGCATTTGGCCAGAGGACCGGGCTGTGCTGGAGACGCTGCCGAGCTCACCTT 1784

DB 781 CCACGTGCACCTTTGGCCAGAGGACCGGCTGTGCTGGAGACGCTGCCGAGCTCACCTT 840
QY 1785 CTTGCTGGAGCTCTTGGGTACCTCTGTTTCTTAGTCACTACAGTGGCTGTGTCTG 1844
DB 841 CTTGCTGGAGCTCTTGGGTACCTCTGTTTCTTAGTCACTACAGTGGCTGTGTCTG 900
QY 1845 GGTGCCAGGCGCGCTCG---CCAGCATCTCTCATCTCACTCAACATGTTCTCTCTT 1901
DB 901 GGTGCCAGGCGCGCTCGGCCCCAGCATCTCTCATCTCACTCAACATGTTCTCTCTT 960
QY 1902 TTCCACAGCCCCAGCAACAGGCTGCTTACCCCCCGGAGGAGGTGTTCCAGGCCACGCT 1961
DB 961 TTCCACAGCCCCAGCAACAGGCTGCTTACCCCCCGGAGGAGGTGTTCCAGGCCACGCT 1020
QY 1962 GGTGCTCTGGCTTTGGCCATGGTCCCATCTCTGCTGCTTGGCACACCCCTGCACCTGCT 2021
DB 1021 GGTGCTCTGGCTTTGGCCATGGTCCCATCTCTGCTGCTTGGCACACCCCTGCACCTGCT 1080
QY 2022 GCACCGCCACCGCCCGCCCTGCGGAGAGGCCCGCTGACCGACAGGAGGAAAAAAGGC 2081
DB 1081 GCACCGCCACCGCCCGCCCTGCGGAGAGGCCCGCTGACCGACAGGAGGAAAAAAGGC 1140
QY 2082 GGGGTTGTGACCTGCTGACGCTCTGTGAATGGCTGGAGCTCCGATGAGGAAAAAGGC 2141
DB 1141 GGGGTTGTGACCTGCTGACGCTCTGTGAATGGCTGGAGCTCCGATGAGGAAAAAGGC 1200
QY 2142 AGGGGGCTTGATGATGAGAGGAGGCCGAGCTGCTCCCTCCGAGGTGCTCATGACCA 2201
DB 1201 AGGGGGCTTGATGATGAGAGGAGGCCGAGCTGCTCCCTCCGAGGTGCTCATGACCA 1260
QY 2202 GGCATCACACCATCTGAGTCTGCTGGGCTGCTGCTCCAAACACGCTCTCTACCTGG 2261
DB 1261 GGCATCACACCATCTGAGTCTGCTGGGCTGCTGCTCCAAACACGCTCTCTACCTGG 1320
QY 2262 CTTGCGGCTTGAAGCTTGGCCACCGCCAGCTGTCGAGGTTCTGTGGGCCATGGTGTAT 2321
DB 1321 CTTGCGGCTTGAAGCTTGGCCACCGCCAGCTGTCGAGGTTCTGTGGGCCATGGTGTAT 1380
QY 2322 GGCATAGGCTTGGGCTGGGCGGAGGTGGGCGGTGGGCGGTGGTGTGGTCCCAT 2381
DB 1381 GGCATAGGCTTGGGCTGGGCGGAGGTGGGCGGTGGGCGGTGGTGTGGTCCCAT 1440
QY 2382 CTTTCCGCTTGGCGGTGATGACCGTGGCTATCTCTGCTGGTGTGAGGAGGACTCTCAGC 2441
DB 1441 CTTTCCGCTTGGCGGTGATGACCGTGGCTATCTCTGCTGGTGTGAGGAGGACTCTCAGC 1500
QY 2442 CTTCTGCAACGCTTGGCGGTGCACTGGGTGGAATTCAGAAAGTTCTACTCAGGCAC 2501
DB 1501 CTTCTGCAACGCTTGGCGGTGCACTGGGTGGAATTCAGAAAGTTCTACTCAGGCAC 1560
QY 2502 GGGCTACAGCTGAGTCCCTTACCTTGGCTGGCCACAGATGACTAGGCGCCACTGCGGT 2561
DB 1561 GGGCTACAGCTGAGTCCCTTACCTTGGCTGGCCACAGATGACTAGGCGCCACTGCGGT 1620
QY 2562 CTTGCGACACCTCTCTTCTGACCTCTGAGGCGAGGA 2596
DB 1621 CTTGCGACACCTCTCTTCTGACCTCTGAGGCGAGGA 1655

RESULT 7
AK004094

LOCUS

DEFINITION

AK004094 1537 bp mRNA linear HTC 02-SEP-2005
Mus musculus 18-day embryo whole body cDNA, RIKEN full-length
enriched library, clone:1110034K16 product:ATPASE, H+ TRANSPORTING,
LYSOSOMAL I (FRAGMENT), full insert sequence.

ACCESSION

AK004094.1 GI:12835141

VERSION

HTC; CAP trapper.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

| | | | |
|----|------|---|------|
| Db | 421 | CTGTGATGTTGGGGATGTGGCCACGGGCTGCTCATGTTCTCTTCGCCCTGGCCATG | 480 |
| Qy | 1318 | GTCCTTGGGAGAACCGACCGGCTGTGAAGCCCGCGAGAACGAGATCTGGCAGACTTTC | 1377 |
| Db | 481 | GTCCTTGGGAGAACCGACCGGCTGTGAAGCCCGCGAGAACGAGATCTGGCAGACTTTC | 540 |
| Qy | 1378 | TTCAGGGGCGCTACCTGCTCTCTGCTTATGGGCTGTTCCTCACTACACGGGCTTCATC | 1437 |
| Db | 541 | TTCAGGGGCGCTACCTGCTCTCTGCTTATGGGCTGTTCCTCACTACACGGGCTTCATC | 600 |
| Qy | 1438 | TACAAACGAGTGTTCAGTCGCGGCCACAGCATCTTCCCTCTGGGCTGGAGTGGCGGCC | 1497 |
| Db | 601 | TACAAACGAGTGTTCAGTCGCGGCCACAGCATCTTCCCTCTGGGCTGGAGTGGCGGCC | 660 |
| Qy | 1498 | ATGGCCAAACAGTCTGGCTGGAGTGTGATCTCTCTGGCCCGCAGCACGATGCTTACCCGTG | 1557 |
| Db | 661 | ATGGCCAAACAGTCTGGCTGGAGTGTGATCTCTCTGGCCCGCAGCACGATGCTTACCCGTG | 720 |
| Qy | 1558 | GATCCCAACGTCACCGGTGTCTTCTGGGACCCTTACCCCTTTGGCATCGATCTATTGG | 1617 |
| Db | 721 | GATCCCAACGTCACCGGTGTCTTCTGGGACCCTTACCCCTTTGGCATCGATCTATTGG | 780 |
| Qy | 1618 | AGCTGTGCTGCCAACCACTTGAGCTTCTCAACTCCTTCAAGATGAAGATGTCCGTGATC | 1677 |
| Db | 781 | AGCTGTGCTGCCAACCACTTGAGCTTCTCAACTCCTTCAAGATGAAGATGTCCGTGATC | 840 |
| Qy | 1678 | CT- GGGGCGTGTGCACATGGCCCTTT- GGGGTGGTCTCTGGAGTCTTCAAACACGTGCACT | 1735 |
| Db | 841 | CTGGGGCGTGTGCACATGGCCCTTTGGGGGTGGTCTCTCGAGTCTTCAAACAGTGCAT | 900 |
| Qy | 1736 | TTGG--CCAGAGGACCGGCTGTCTGTGGAGACGCTCCCGGAGCTCACTTCCTGCTGGG | 1793 |
| Db | 901 | TTGGGGCAGAAAGACCSGGCTGTCTGTGGAGACGCTCCCGGAGCTCACTCTNGCTGGG | 960 |
| Qy | 1794 | ACTCTTCGGTTCCTCGTGTTCCTAGTGCATCTACAAGTGGCTGTCTCTGGGCTGCCAG | 1853 |
| Db | 961 | ACTCTTCGGTTCCTCGTGTTCCTAGTGCATCTACAAGTGGGTGTGTGTGGTGCAGG | 1020 |
| Qy | 1854 | GGCGGCTCTGCCGAGCATCTCATFCACTTCATCAACATGTTCTCTTTC | 1902 |
| Db | 1021 | GGCGCTCGGCCCAACATCTCTCATCTMMTTCATAAATGTTCYTTVTTTC | 1069 |

| | |
|------------|--|
| RESULT 9 | |
| BX358008 | |
| LOCUS | 909 bp mRNA linear EST 08-APR-2004 |
| DEFINITION | BX358008 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CSODI032YN10 5-PRIME, mRNA sequence. |
| ACCESSION | BX358008 |
| VERSION | BX358008.2 GI:46285730 |
| KEYWORDS | EST. |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo. |
| REFERENCE | 1 (bases 1 to 909) |
| AUTHORS | Li, W.B., Gruber, C., Jesse, J. and Polayes, D. |
| TITLE | Full-length cDNA libraries and normalization |
| JOURNAL | Unpublished (2001) |
| COMMENT | On May 5, 2003 this sequence version replaced gi:30366180. Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6186.f For more information about this cluster, see |

http://www.genoscope.cns.fr/cdna?s=CS0DI032DG050P1&c=6186.f.

| FEATURES | source | Location/Qualifiers |
|-----------------------|-------------------|---|
| | 1..909 | |
| | | /organism="Homo sapiens" |
| | | /mol_type="mRNA" |
| | | /db_xref="taxon:9606" |
| | | /clone="CS0DI032YN10" |
| | | /tissue_type="PLACENTA COT 25-NORMALIZED" |
| | | /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED" |
| | | /notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized." |
| ORIGIN | | |
| Query Match | 33.5%; | Score 885.4; DB 4; Length 909; |
| Best Local Similarity | 99.3%; | Pred. No. 4.2e-161; |
| Matches | 889; Conservative | 0; Mismatches 6; Indels 0; Gaps 0; |
| Qy | 599 | CCCTAGAGCGCTGCTCTGGAGGGCTGCGGGGCTTCTCTCATATGCGACATTCAGGAGC 658 |
| Db | 1 | CCCGGATCGCTGCTCTGGAGGGCTGCGGGGCTTCTCTCATATGCGACATTCAGGAGC 60 |
| Qy | 659 | TGAGCAGCGCTGGAGCACCCGCTACGCGGAGCAGCCACGCTGGATGACCTTCTCA 718 |
| Db | 61 | TGGAGCAGCGCTGGAGCACCCGCTGTCGCGGAGCAGCCACGCTGGATGACCTTCTCA 120 |
| Qy | 719 | TCTCTACTTGGGTGAGCAGATCGGACAGAGATCGCAAGATCACGGACTGCTTCCACT 778 |
| Db | 121 | TCTCTACTTGGGTGAGCAGATCGGACAGAGATCGCAAGATCACGGACTGCTTCCACT 180 |
| Qy | 779 | GCCACGCTTCCCGTTTCTGCAGCAGGAGGAGCCCGCTCGGGGCCCTGCAGCAGCTGC 838 |
| Db | 181 | GCCACGCTTCCCGTTTCTGCAGCAGGAGGAGCCCGCTCGGGGCCCTGCAGCAGCTGC 240 |
| Qy | 839 | AACAGCAGACCGAGAGCTGCAGAGGTCTCTGGGAGACAGAGCGGTTCTTGAGCCAGG 898 |
| Db | 241 | AACAGCAGACCGAGAGCTGCAGAGGTCTCTGGGAGACAGAGCGGTTCTTGAGCCAGG 300 |
| Qy | 899 | TGCTAGCGGGGTGCTGCAGCTGCTGCGCGCAGGCGAGTGAGGTCACAAGATGAAGG 958 |
| Db | 301 | TGCTAGCGGGGTGCTGCAGCTGCTGCGCGCAGGCGAGTGAGGTCACAAGATGAAGG 360 |
| Qy | 959 | CGGTGTACCTGGCCCTTGAAACCACTGCAGCGTGAGCACACGACGAAGTGCCCTCATTCGCCG 1018 |
| Db | 361 | CGGTGTACCTGGCCCTTGAAACCACTGCAGCGTGAGCACACGACGAAGTGCCCTCATTCGCCG 420 |
| Qy | 1019 | AGGCTGTGTCTGTGCGAGACCTTGCCCGCTCTGAGAGGCGCTTCGGGACAGCTCA 1078 |
| Db | 421 | AGGCTGTGTCTGTGCGAGACCTTGCCCGCTCTGAGAGGCGCTTCGGGACAGCTCA 480 |
| Qy | 1079 | TGGAGGAGGAGTGAGTGGCGGTGCTACCGCATCCCTGCGGAGCATGCCCGCCACAC 1138 |
| Db | 481 | TGGAGGAGGAGTGAGTGGCGGTGCTACCGCATCCCTGCGGAGCATGCCCGCCACAC 540 |
| Qy | 1139 | TGATCGCACCAACCGCTTTCAGGGCAGCTTCAGGGCATCGTGATCGCTACGGCGTGG 1198 |
| Db | 541 | TGATCGCACCAACCGCTTTCAGGGCAGCTTCAGGGCATCGTGATCGCTACGGCGTGG 600 |
| Qy | 1199 | GCGCTTACAGAGGTCAACCCCGCTTCAACCATCATACCTTCCCTTCTGTGTTG 1258 |
| Db | 601 | GCGCTTACAGAGGTCAACCCCGCTTCAACCATCATACCTTCCCTTCTGTGTTG 660 |
| Qy | 1259 | CTGTGATGTTCGGGATGTGGGCCAGGGCTGCTCATGTTCTCTTTCGCCCTGGCATGG 1318 |
| Db | 661 | CTGTGATGTTCGGGATGTGGGCCAGGGCTGCTCATGTTCTCTTTCGCCCTGGCATGG 720 |
| Qy | 1319 | TCTTTCGAGAACACGACCGGTGTGAAGCGCGCAGAACAGATCTGGCAGACTTCT 1378 |
| Db | 721 | TCTTTCGAGAACACGACCGGTGTGAAGCGCGCAGAACAGATCTGGCAGACTTCT 780 |
| Qy | 1379 | TGAGGGCGCGCTACCTGCTTCTGCTTTATGGGCGCTGTTCTTCCATCTACACCGGCTCATCT 1438 |

Db 781 TCAGGGGCGCTACCTGCTCTCTGCTTATGGGCGCTGTTCTCCATCTACACCGGCTTCATCT 840
QY 1439 ACAACGAGTCTTCACTGCGGCCACACAGCATCTTCCCTCGGCTGGAGTGGC 1493
Db 841 ACAACGAGTCTTCACTGCGGCCACACAGCATCTTCCCTCGGCTGGAGTGGC 895

RESULT 10

BX385317/c
LOCUS BX385317 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DL009YP09 3-PRIME, mRNA sequence.

ACCESSION

VERSION BX385317
KEYWORDS BX385317.2 GI:46575483
SOURCE EST.

ORGANISM

Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE

1. (bases 1 to 1067)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)

JOURNAL

On May 8, 2003 this sequence version replaced gi:30451332.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

AUTHORS

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6186.f

FEATURES

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?e=CS0DL009CH05NP1&c=6186.f.

Location/Qualifiers

1..1067
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL009YP09"
/cell_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 33.2%; Score 877.4; DB 4; Length 1067;
Best Local Similarity 96.4%; Pred. No. 1.5e-159;
Matches 942; Conservative 17; Mismatches 10; Indels 8; Gaps 6;
QY 1580 TCCTGGGACCTACCCCTTGGCATCGATCCTATTGGAGCCTGGCTGCCAACCACTTGA 1639
Db 1009 TTCTGGGACCTACCCCTTGGCATCGATCCTATTGGAGCCTGGCTGCCAACCACTTGA 950
QY 1640 GCTTCCTCAACTCCTT-CAAGATGAAGATGTCCTGTCATCC-TGGGCGTGTGCATGGC 1697
Db 949 GCTTCCTCAACTCCTTCCAGATGAAGATGTCCTGTCATCCYTTGGGCGTGTGCATGGC 890
QY 1698 CTTTGGGCGTGTCTCGGAGTCTTCAACCACTGACCTTTGGCCAGAGCACCGGTGCT 1757
Db 889 CTTTGGGCGTGTCTCGGAGTCTTCAACCACTGACCTTTGGCCAGAGCACCGGTGCT 830
QY 1758 GCTGGAGACGCTGCCGGAGCTCACCTTCTGCTGGAGCTCTTCGGTTACCTGCTGCTCT 1817
Db 829 GCTGGAGACGCTGCCGGAGCTCACCTTCTGCTGGAGCTCTTCGGTTACCTGCTGCTCT 770

QY 1818 AGTCATCTACAAAGTGGCTGTGTGCTGGGCTGCCAGGCGCGCTCG-CCAGCATCCT 1874
Db 769 AGTCATCTACAAAGTGGCTGTGTGCTGGGCTGCCAGGCGCGCTCGCCCCAGCATCCT 710
QY 1875 CATCACTTCAATCAACATGTTCTCTTCTCCACAGCCCCAGCAACAGGCTGTCTTACCC 1934
Db 709 CATCACTTCAATCAACATGTTCTCTTCTCCACAGCCCCAGCAACAGGCTGTCTTACCC 650
QY 1935 CCGGAGAGGTGGTCCAGGCCACAGCTGGTGTCTTGGCTTGGCCATGGTCCCATCCT 1994
Db 649 CCGGAGAGGTGGTCCAGGCCACAGCTGGTGTCTTGGCTTGGCCATGGTGHATCCT 590
QY 1995 GCTGCTTGGCACACCCCTGCACTGCTGCACCGCCACCGCCCGCTCGGAGGAGGCC 2054
Db 589 GCTGCTTGGCACACCCCTGCTGCTGCAACCGCCACCGCCCGG-CTCGGAGGAGGCC 531
QY 2055 CGTCACCGACAGGAGGAAACAAGCGCCGGGTGTGGACCTTGCCTGACGCATCTGTGAA 2114
Db 530 CGTCACCGACAGGAGGAAACAAGCGCCGGGTGTGCTGACCTTCCCTGACGCATCTGTGAA 471
QY 2115 TGGCTGGAGCTCCGATGAGGAAAGAGGAGGCGCTTGGATGATGAAGAGGAGGCCAGCT 2174
Db 470 TGGCTGGAGCTCCGATGAGGAAAGAGGAGGCGCTTGGATGATGAAGAGGAGGCCAGCT 411
QY 2175 CGTCCCTCCGAGGTGCTCATGCACAGGCCATCCACACCATCGAGTTCTGCTGGGCTG 2234
Db 410 CGTCCCTCCGAGGTGCTCATGCACAGGCCATCCACACCATCGAGTTCTGCTGGGCTG 351
QY 2235 CGTCTCCAACACCGCTCTCTACTCGGCTGTGGGCCCTTGAGCCCTGGGCCACGCCAGCT 2294
Db 350 CGTCTCCAACACCGCTCTCTACTCGGCTGTGGGCCCTTGAGCCCTGGGCCACGCCAGCT 291
QY 2295 GTCCAGAGTCTTGTGGGCCATGGTATGCGCATAGGCTTGGGCTGGGCCGGGAGGTGG 2354
Db 290 GTCCAGAGTCTTGTGGGCCATGGTATGCGCATAGGCTTGGGCTGGGCCGGGAGGTGGT 231
QY 2355 CGTGGCGGCTGTGGTGGTCCCATCTTTGCGCCCTTTGCGCGTGTGATGACCGTGGCTAT 2414
Db 230 CGTGGCKGTTGTGGTGGT-GTCCCATCTTTGCGCGCTTTGCGCGTGTGATGACCGTGGCTAT 172
QY 2415 CTTGCTGTGTGATGGAGGACTTCTAGCCTTCTGACGCTTCTGACGCGCTTGGCTGCACTGGGTGA 2474
Db 171 CTTGCTGTGTGATGGAGGACTTCTAGCCTTCTGACG-CTTGGCGCTGCACTGGGTGA 113
QY 2475 ATTCCAGAACAGTTCTACTCAGGCACGGCTTACAGCTGAGTCCCTTCACTTCCCTGCTGC 2534
Db 112 ATTCCAGAACAGTTCTACTCAGGCACGGGCTACAGGCTGAGTCCCTTCACTTCCCTGCTGC 53
QY 2535 CACAGATGACTAGGCGC 2551
Db 52 CACAGATGACTAGGSSS 36

RESULT 11

LOCUS BQ950273
DEFINITION BQ950273 941 bp mRNA linear EST 21-AUG-2002
5', mRNA sequence.
AGENCOURT_8835676 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:6421778
ACCESSION BQ950273
VERSION BQ950273.1 GI:22365751
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1. (bases 1 to 941)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar

cdna Library Preparation: Rubin Laboratory
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LUCM2597 row: 1 column: 03
High quality sequence stop: 670.

Location/Qualifiers
1. .941

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6421778"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 18"
/notes="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cdna made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN

Query Match 33.0%; Score 870; DB 3; Length 941;
Best Local Similarity 98.8%; Pred. No. 4e-158;
Matches 928; Conservative 0; Mismatches 6; Indels 5; Gaps 5;

```
QY 647 GCTTCAGGAGCTGGAGCAGCGCTGGAGCACCCTCGTACGGCGGAGCCAGCCAGCTGGA 706
DB 1 GCTTCAGGAGCTGGAGCAGCGCTGGAGCACCCTCGTACGGCGGAGCCAGCCAGCTGGA 60
QY 707 TGACCTTCCTCATCTCTCTACTGGGGTGAGCAGATCGGACAGAGATCCGCAAGATCACGG 766
DB 61 TGACCTTCCTCATCTCTCTACTGGGGTGAGCAGATCGGACAGAGATCCGCAAGATCACGG 120
QY 767 ACTGCTTCCACTGCCAGCTTCTCCGTTTCTGCAGCAGGAGGAGCCGCTCGGGGCC 826
DB 121 ACTGCTTCCACTGCCAGCTTCTCCGTTTCTGCAGCAGGAGGAGGAGCCGCTCGGGGCC 180
QY 827 TGACAGAGCTGCAACAGCAGCAGCAGAGCTGCAGAGGTCCTCGGGGAGACAGAGCGT 886
DB 181 TGACAGAGCTGCAACAGCAGCAGCAGAGCTGCAGAGGTCCTCGGGGAGACAGAGCGT 240
QY 887 TCCTGAGCAGAGCTAGGCGGGTGCTGAGCTGCTGCGCCAGGAGGAGGTCAGGTTC 946
DB 241 TCCTGAGCAGAGCTAGGCGGGTGCTGAGCTGCTGCGCCAGGAGGAGGTCAGGTTC 300
QY 947 ACAAGATGAAGGCGCTGTACTCGGCTCTGAACAGTGCAGCGTGAGCACCACGCAAGT 1006
DB 301 ACAAGATGAAGGCGCTGTACTCGGCTCTGAACAGTGCAGCGTGAGCACCACGCAAGT 360
QY 1007 GCCTCATATGCGAGGCTGTGTCTGTGCGAGACTGTCGCCGCTCGAGAGGCGCTGC 1066
DB 361 GCCTCATATGCGAGGCTGTGTCTGTGCGAGACTGTCGCCGCTCGAGAGGCGCTGC 420
QY 1067 GGACAGCTCGATGGAGGAGGAGTGAGTGCCTGCTCAGCCGATCCCTCGCGGACA 1126
DB 421 GGACAGCTCGATGGAGGAGGAGTGAGTGCCTGCTCAGCCGATCCCTCGCGGACA 480
QY 1127 TGCCCCCACCACATCATCCGACCAACCGCTTTCACGGCCAGCTTCCAGGGGATCGTGATC 1186
DB 481 TGCCCCCACCACATCATCCGACCAACCGCTTTCACGGCCAGCTTCCAGGGGATCGTGATC 540
QY 1187 GCTACGGCTGGGCGCTACACAGAGGTCAACCCCGCTCCCTACACATCATCATCTTC 1246
DB 541 CTTACGGCTGGGCGCTACACAGAGGTCAACCCCGCTCCCTACACATCATCATCTTC 600
QY 1247 CTTTCCTGTTGCTGTGATGTTTCGGGGATGTGGGCCACGGGCTGCTCATGTTCTCTTCG 1306
```

```
DB 601 CTTTCCTGTTGCTGTGATGTTTCGGGATGTGGGCACAGGGCTGCTCATGTTCTCTTCG 660
QY 1307 CCCTGGCCATGGTCTTTCGGGAG-AACCGACCGGCTGTGAAAGCCGCGCAGAACGAGATC 1365
DB 661 CCCTGGCCATGGTCTTTCGGGAGAACCGACCGGCTGTGAAAGCCGCGCAGAACGAGATC 720
QY 1366 TGGCAGAGCTTTCCTCAGGGGCGCTACCTGCTCTCTGCTTATGGGCTGTTCCTCATCTAC 1425
DB 721 TGGCAGAGCTTTCCTCAGGGGCGCTACCTGCTCTCTGCTTATGGGCTGTTCCTCATCTAC 780
QY 1426 ACCGGCTTTCATCTACACAGAGTCTTCAGTCGCGCCACACAGCATCTTCCCT-CGGGCTG 1484
DB 781 ACCGGCTTTCATCTACACAGAGTCTTCAGTCGCGCCACACAGCATCTTCCCTCGGGCTG 840
QY 1485 GAGTGTGGCGCCATCGCCAAACAGTCTGGCTGGAGTGATGCA-TTCTGGCCCGCAGCACA 1543
DB 841 GAGTGTGGCGCCATCGCCAAACAGTCTGGCTGGAGTGATGCA-TTCTGGCCCGCAGCACA 900
QY 1544 CG-ATGCTTACCTGG-ATCCCAACGTCACCGGTGCTT 1580
DB 901 CGAATGCTTACCTGGNATCCCAACGTCACCGGTGCTT 939
```

RESULT 12

AL575554/c
LOCUS AL575554 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
DEFINITION clone CS0DI059YF06 3-PRIME, mRNA sequence.

ACCESSION AL575554
VERSION AL575554.3 GI:46248511
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 977)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 16, 2001 this sequence version replaced gi:31313861.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6186.f

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?e=CS0DI059DC03NP1&c=6186.f.

FEATURES

source

1. .977
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI059YF06"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 32.9%; Score 867.8; DB 1; Length 977;
Best Local Similarity 95.1%; Pred. No. 1.1e-157;
Matches 915; Conservative 18; Mismatches 23; Indels 6; Gaps 4;
QY 1623 GGTGCGCAACCACTTGAGCTTCTCACTTCAAGATGAAGATCGCTCATCTCTGGG 1682

```
Db 972 GCCTGCTCCMACCACTTGRSTTCTCCTCAMCCCTTCAAGATRAAR-TGTCCGTATCATCTCGS 914
Qy 1683 COTCTGTGCACATGGCTTTGGGGTGTCTCTCGAGTCTTCAACCAAGTCACTTTTGCCCA 1742
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 913 -STCGTGCAATGGCTTTGGGGTGTCTCTCGAGTCTTCAACCAAGTCACTTTTGCCCA 855
Qy 1743 GAGGACACCGCTGTCTGTGAGACGCTGCGGAGCTCACTTCTCTGTCTGTGGACTCTTGG 1802
Db 854 GAGGACACCGCTGTCTGTGAGACGCTGCGGAGCTCACTTCTCTGTCTGTGGACTCTTGG 795
Qy 1803 TTACTCTGTCTTCTAGTCACTTACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1862
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 794 TTMCCCTGTCTTCTAGTCACTTACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 735
Qy 1863 G---CCAGCATCTCATCACTTATCAACATGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1919
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 734 GGGCCCCAGCATCTCATCACTTATCAACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 675
Qy 1920 CAGGCTGTCTTACCCCGGACGAGGTGTCCAGGACGCTGTGTGTGTGTGTGTGTGTGTGTGT 1979
Db 674 CAGGCTGTCTTACCCCGGACGAGGTGTCCAGGACGCTGTGTGTGTGTGTGTGTGTGTGTGT 615
Qy 1980 CATGTGTGCTTCTGTCTTGTGCAACCTTGTGCACTGTGTGCACTGTGTGCACTGTGTGCA 2039
Db 614 CATGTGTGCTTCTGTCTTGTGCAACCTTGTGCACTGTGTGCACTGTGTGCACTGTGTGCA 555
Qy 2040 CTTGGGAGGAGCGGCTGTGCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2099
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 554 -CTGCGGAGNATGCCCGCTTACCNACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 496
Qy 2100 TGACGCATCTGTGAATGGCTGTGAGCTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2159
Db 495 TGACGCATCTGTGAATGGCTGTGAGCTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 436
Qy 2160 AGAGAGGCGGAGCTGTCTTCCCTCCGAGTGTCTATGCAACAGGAGGAGGAGGAGGAGGAGGAG 2219
Db 435 AGAGAGGCGGAGCTGTCTTCCCTCCGAGTGTCTATGCAACAGGAGGAGGAGGAGGAGGAGGAG 376
Qy 2220 GTTCTGCTGGCTGGCTGTCTCAACACCGCTTCTACCTGCGCTGTGTGTGTGTGTGTGTGTGT 2279
Db 375 GTTCTGCTGGCTGGCTGTCTCAACACCGCTTCTACCTGCGCTGTGTGTGTGTGTGTGTGTGT 316
Qy 2280 GGGCCACGCGGAGCTGTGCGAGTGTCTGTGGCCATGTGTGATGCGATAGGCTGTGGCT 2339
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 315 GGGCCACGCGGAGCTGTGCGAGTGTCTGTGGCCATGTGTGATGCGATAGGCTGTGGCT 256
Qy 2340 GGGCCGAGGAGTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2399
Db 255 GGGCCGAGGAGTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 196
Qy 2400 GATGACGCTGCTATCTCTGCTGTGATGAGGAGGAGTCTTCTGAGCTTCTGAGCGCTTGG 2459
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 195 GATGACGCTGCTATCTCTGCTGTGATGAGGAGGAGTCTTCTGAGCTTCTGAGCGCTTGG 136
Qy 2460 GCTGACCTGGTGGATTTCCAGAACAAAGTTCTACTCAGGACGCGCTTCAAGCTGAGTCC 2519
Db 135 GCTGACCTGGTGGATTTCCAGAACAAAGTTCTACTCAGGACGCGCTTCAAGCTGAGTCC 76
Qy 2520 TTTCACCTTGTGTGCAAGATGATAGGCGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGT 2579
Db 75 TTTCACCTTGTGTGCAAGATGATAGGCGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGT 16
Qy 2580 TG 2581
Db 15 TG 14
```

```
RESULT 13
BX385318 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
LOCUS
DEFINITION
BX385318 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cdna clone CS0DL009YP09 5-PRIME, mRNA sequence.
ACCESSION
BX385318
```

```
VERSION BX385318.2 GI:46620145
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE 1 (bases 1 to 1074)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30453326.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, Cp 5706 - 91057 EVRY cedex - FRANCE
Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6186.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DL009CH05QP1&c=6186.f.
Location/Qualifiers
1..1074
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL009YP09"
/cell_type="B CELLS (RAMOS CELL LINE)"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 32.7%; Score 863.4; DB 4; Length 1074;
Best Local Similarity 97.3%; Pred. No. 7.8e-157;
Matches 923; Conservative 6; Mismatches 14; Indels 6; Gaps 5;
Qy 3 GCGTCGCGGAGCGGCGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 61
Db 48 GTGCGCGGAGGNGNGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 107
Qy 62 GCTCCATGTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 121
Db 108 GCTCCATGTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 167
Qy 122 CTTACACCTGCTGAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 181
Db 168 CTTACACCTGCTGAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 227
Qy 182 CTTACACCTGCTGAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 241
Db 228 CTTACACCTGCTGAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 287
Qy 242 AGAAGACCTTCACTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 301
Db 288 AGAAGACCTTCACTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 347
Qy 302 CAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 361
Db 348 CAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 407
Qy 362 AGCGCTGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 421
Db 408 AGCGCTGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 466
Qy 422 TGCACACGCTGAGCTTCCACGCGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 481
```


BX357597/c
LOCUS BX357597 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1028Y106 3-PRIME, mRNA sequence.
ACCESSION BX357597
VERSION BX357597.2 GI:46287556
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 921)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 5, 2003 this sequence version replaced gi:30368137.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6186.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0D1028Y106&NP1&c=6186.f.

FEATURES

Location/Qualifiers
1..921
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1028Y106"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 32.2%; Score 850.2; DB 4; Length 921;
Best Local Similarity 98.7%; Pred. No. 2.8e-154;
Matches 895; Conservative 4; Mismatches 2; Indels 6; Gaps 4;
QY 1678 CTGGCGCTCGTGCAATGCGCTTTGGGGTGGTCTCGGAGTCTTCAACACCGTGCACCTTT 1737
DB 910 CTGGCGCTCGTGCAATGCGCTTTGGGGTGGTCTCGGAGTCTTMAACACCGTGCACCTTT 852
QY 1738 GGCACAGGACCGGCTGCTGCGAGCGCTGCGGAGCTCACCTTCTGCTGGGACTC 1797
DB 851 GGCACAGGACCGGCTGCTGCGAGCGCTGCGGAGCTCACCTTCTGCTGGGACTC 792
QY 1798 TTGCGTTACTGCTGCTTCTAGTCTCATCAAGTGGCTGTGTCTGGGCTGCCAGGGCC 1857
DB 791 TTGCGTTACTGCTGCTTCTAGTCTCATCAAGTGGCTGTGTCTGGGCTGCCAGGGCC 732
QY 1858 GCCTCG---CCGAGCATCTCATCCACTTTCATCAAGTTCCTCTTCTCCACAGCCCC 1914
DB 731 GCCTCGGCCCCGAGCATCTCATCAAGTTCCTCTTCTCCACAGCCCC 672
QY 1915 AGCAACAGGCTGCTTACCCCGGAGGAGTGGTCCAGGCCACCGTGGTGGTCTGGCC 1974
DB 671 AGCAACAGGCTGCTTACCCCGGAGGAGTGGTCCAGGCCACCGTGGTGGTCTGGCC 612
QY 1975 TTGGCCATGGTCCCATCTGCTGTGGCACACCCCTGCACCTGCTGCACCGCCACCGC 2034
DB 611 TTGGCCATGGTGGCCATCTGCTGTGGCACACCCCTGCACCTGCTGCACCGCCACCGC 552
QY 2035 CGCCCGCTGCGGAGGAGGCGCCCTGACCGACAGGAGGAGAAACAAGGCGGGTTGCTGGAC 2094

DB 551 CGCGCGCTCGGAAGGAGGCGCTGACCGACAGGAGGAGAAACAAGCGCGGGTTGCTGGAC 492
QY 2095 CTGGCTGACGCATCTGTGAATGGCTGGAGCTCCGATGAGGAGAAAGCAGCGGGCCCTGGAT 2154
DB 491 CTGGCTGACGCATCTGTGAATGGCTGGAGCTCCGATGAGGAGAAAGCAGCGGGCCCTGGAT 432
QY 2155 GATGAAGAGGAGGCGGAGCTGCTGCCCTCCGAGGTGCTCATGCACAGGCGCATCCACACC 2214
DB 431 GATGAAGAGGAGGCGGAGCTGCTGCCCTCCGAGGTGCTCATGCACCA-GCCATCCACACC 373
QY 2215 ATCGAGTTCTGCTGGCTGGCTGCTCTCCAAACACCGCTCTACCTGCGCCTGTGGGGCCCTG 2274
DB 372 ATCGAGTTCTGCTGGCTGGCTGCTCTCCAAACACCGCTCTACCTGCGCCTGTGGGGCCCTG 313
QY 2275 AGCCTGGCCCCACGCCAGCTGTCG-AGTTTCTGTGGGCCATGGTATGGGCATAGGCCT 2333
DB 312 AGCCTGGCCCCACGCCAGCTGTCGAAAGGTTCTGTGGGCCATGGTATGGCATAGGCCT 253
QY 2334 GGGCCTGGGCCCGGAGGTGGCGCTGTGGTGTGGTCCCATCTTTTGGCGCCCTT 2393
DB 252 GGGCCTGGGCCCGGAGGTGGCGCTGTGGTGTGGTCCCATCTTTTGGCGCCCTT 193
QY 2394 TGCCGTGATGACCGTGGCTATCTCTGTGTGATGGAGGACTCTCAGCCTTCTGTCACGC 2453
DB 192 TGCCGTGATGACCGTGGCTATCTCTGTGTGATGGAGGACTCTCAGCCTTCTGTCACGC 133
QY 2454 CTGGCGGCTGCACTGGGTGGAAATTCAGAAAGTTCTTACTCAGGACCGGGCTACAGCT 2513
DB 132 CTGGCGGCTGCACTGGGTGGAAATTCAGAAAGTTCTTACTCAGGACCGGGCTACAGCT 73
QY 2514 GAGTCCCTTACCTTTCGCTGCCACAGATGACTAGGGGCCACCTGCAAGTCTTCCAGACCT 2573
DB 72 GAGTCCCTTACCTTTCGCTGCCACAGATGACTAGGGGCCACCTGCAAGTCTTCCAGACCT 13
QY 2574 CTTTCT 2580
DB 12 CTTTCT 6

Search completed: June 30, 2006, 07:33:48
Job time : 11162 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2006, 05:18:12 ; Search time 343 Seconds
(without alignments)
9101.427 Million cell updates/sec

Title: US-10-783-519-1
Perfect score: 2640
Sequence: 1 cggcgfcgagcgagcgagcag.....gcaaaaaaaaaaaaaa 2640

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 809770 seqs, 591248006 residues

Total number of hits satisfying chosen parameters: 1619540

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA New.*
1: /EMC Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB.seq.*
2: /EMC Celerra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /EMC Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /EMC Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB.seq.*
5: /EMC Celerra_SIDS3/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
6: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB.seq.*
7: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /EMC Celerra_SIDS3/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-----------------------|
| 1 | 2592.2 | 98.2 | 2700 | 6 | US-10-511-937-410 |
| 2 | 1474.2 | 55.8 | 1696 | 7 | US-11-266-748A-76629 |
| 3 | 1474.2 | 55.8 | 1696 | 7 | US-11-266-748A-109489 |
| 4 | 1474.2 | 55.8 | 1696 | 7 | US-11-266-748A-129440 |
| 5 | 1093.2 | 41.4 | 1839 | 7 | US-11-266-748A-361154 |
| 6 | 1093.2 | 41.4 | 1839 | 7 | US-11-266-748A-444533 |
| 7 | 973.2 | 36.9 | 1000 | 7 | US-11-266-748A-221798 |
| 8 | 973.2 | 36.9 | 1000 | 7 | US-11-266-748A-287153 |
| 9 | 973.2 | 36.9 | 1000 | 7 | US-11-266-748A-338582 |
| 10 | 973.2 | 36.9 | 1000 | 7 | US-11-266-748A-397768 |
| 11 | 973.2 | 36.9 | 1000 | 7 | US-11-266-748A-468814 |
| 12 | 707.2 | 26.8 | 738 | 7 | US-11-266-748A-76628 |
| 13 | 707.2 | 26.8 | 738 | 7 | US-11-266-748A-109488 |
| 14 | 707.2 | 26.8 | 738 | 7 | US-11-266-748A-129439 |
| 15 | 674.4 | 25.5 | 1042 | 7 | US-11-266-748A-76627 |
| 16 | 674.4 | 25.5 | 1042 | 7 | US-11-266-748A-109487 |
| 17 | 674.4 | 25.5 | 1042 | 7 | US-11-266-748A-129438 |
| 18 | 668.6 | 25.3 | 671 | 7 | US-11-266-748A-54130 |
| 19 | 639 | 24.2 | 4681 | 7 | US-11-266-748A-29886 |
| 20 | 617.2 | 23.4 | 3137 | 7 | US-11-266-748A-30495 |
| 21 | 551.2 | 20.9 | 561 | 7 | US-11-266-748A-362824 |
| 22 | 551.2 | 20.9 | 561 | 7 | US-11-266-748A-466203 |
| 23 | 549.4 | 20.8 | 831 | 7 | US-11-266-748A-8067 |
| 24 | 532.8 | 20.2 | 705 | 7 | US-11-266-748A-367398 |
| 25 | 532.8 | 20.2 | 705 | 7 | US-11-266-748A-450777 |

| | | | | | | |
|----|-------|------|------|---|-----------------------|--------------------|
| 26 | 531.6 | 20.1 | 884 | 7 | US-11-266-748A-76631 | Sequence 76631, A |
| 27 | 531.6 | 20.1 | 884 | 7 | US-11-266-748A-109491 | Sequence 109491, A |
| 28 | 531.6 | 20.1 | 884 | 7 | US-11-266-748A-129442 | Sequence 129442, A |
| 29 | 461.6 | 17.5 | 1469 | 7 | US-11-266-748A-98830 | Sequence 98830, A |
| 30 | 461.6 | 17.5 | 1469 | 7 | US-11-266-748A-151641 | Sequence 151641, A |
| 31 | 345 | 13.1 | 1151 | 7 | US-11-266-748A-185488 | Sequence 185488, A |
| 32 | 345 | 13.1 | 1151 | 7 | US-11-266-748A-240948 | Sequence 240948, A |
| 33 | 340.8 | 12.9 | 1127 | 7 | US-11-266-748A-76630 | Sequence 76630, A |
| 34 | 340.8 | 12.9 | 1127 | 7 | US-11-266-748A-109490 | Sequence 109490, A |
| 35 | 340.8 | 12.9 | 1127 | 7 | US-11-266-748A-129441 | Sequence 129441, A |
| 36 | 308.8 | 11.7 | 444 | 6 | US-10-488-619-1101 | Sequence 1101, Ap |
| 37 | 301 | 11.4 | 446 | 6 | US-10-488-619-1100 | Sequence 1100, Ap |
| 38 | 262.4 | 9.9 | 3310 | 6 | US-10-449-902-13157 | Sequence 13157, A |
| 39 | 225 | 8.5 | 1147 | 7 | US-11-266-748A-260923 | Sequence 260923, A |
| 40 | 225 | 8.5 | 1147 | 7 | US-11-266-748A-321440 | Sequence 321440, A |
| 41 | 220 | 8.3 | 1548 | 7 | US-11-266-748A-99200 | Sequence 99200, A |
| 42 | 220 | 8.3 | 1548 | 7 | US-11-266-748A-152011 | Sequence 152011, A |
| 43 | 191.6 | 7.3 | 1000 | 7 | US-11-266-748A-223064 | Sequence 223064, A |
| 44 | 179.4 | 6.8 | 598 | 6 | US-10-488-619-2079 | Sequence 2079, Ap |
| 45 | 166.2 | 6.3 | 605 | 7 | US-11-266-748A-51893 | Sequence 51893, A |

ALIGNMENTS

RESULT 1
US-10-511-937-410
; Sequence 410, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 410
; LENGTH: 2700
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-511-937-410

| | | | | |
|-----------------------|--------------|---------------|--|--------------|
| Query Match | 98.2%; | Score 2592.2; | DB 6; | Length 2700; |
| Best Local Similarity | 99.5%; | Pred. No. 0; | | |
| Matches 2599; | Conservative | 0; | Mismatches | 9; |
| | | | Indels | 3; |
| | | | Gaps | 1; |
| QY | 33 | GC | CGCGCGGAGCAGACACCCGGGGACCATGGGCTCCATGTTCCGAGCGAGGAGTGCCCT | 92 |
| Db | 84 | GC | CGCGCGGAGCAGACACCCGGGGACCATGGGCTCCATGTTCCGAGCGAGGAGTGCCCT | 143 |
| QY | 93 | GGT | CCAGCTCTTCTTCTCCACACAGCGGCTGCTACACTCGTGGTGGCTGGCGGCGCT | 152 |
| Db | 144 | GGT | CCAGCTCTTCTTCTCCACACAGCGGCTGCTACACTCGTGGTGGCTGGCGGCGCT | 203 |
| QY | 153 | GG | CCCTCGTGGAGTTCCAGAGACCTCAACGCTCGGTGAGCGCTTCCAGAGAGCGCTTTGT | 212 |
| Db | 204 | GG | CCCTCGTGGAGTTCCAGAGACCTCAACGCTCGGTGAGCGCTTCCAGAGAGCGCTTTGT | 263 |

213 GGTGATGTTGGCGCTGTGAGAGCTGGAGAACCTTCACTTCTCTGACGAGAGGT 272
264 GGTGATGTTGGCGCTGTGAGAGCTGGAGAACCTTCACTTCTCTGAGAGAGGT 323
273 GGGCGGGCTGGGCTGGTCTTGCCTCCCGCAAAGGGAGGCTGCGGCAACCCCAACCCCG 332
324 GGGCGGGCTGGGCTGGTCTTGCCTCCCGCAAAGGGAGGCTGCGGCAACCCCAACCCCG 383
333 GGAACCTGCTGCGCATCCAGAGGAGAGCGAGCGCTGGGCCCAAGGAGCTGCGGATGTGG 392
384 GGAACCTGCTGCGCATCCAGAGGAGAGCGAGCGCTGGGCCCAAGGAGCTGCGGATGTGG 443
393 GGGCAACCAAGAGGCGCTGCGGCGCCAGCTGCAACAGCTGCAAGCTGCAAGCGCGCTGCT 452
444 GGGCAACCAAGAGGCGCTGCGGCGCCAGCTGCAACAGCTGCAAGCTGCAAGCGCGCTGCT 503
453 AGCCAGGGCCATGAACCTCAGCTGGCAGCGGCCCAACAGATGGGGCTCAGAGAGGAC 512
504 AGCCAGGGCCATGAACCTCAGCTGGCAGCGGCCCAACAGATGGGGCTCAGAGAGGAC 563
513 GCCCTGCTCCAGGCGCCCGGGGGCGGCAACAGGAGCTGAGGGTCACTTTGTGGCAGG 572
564 GCCCTGCTCCAGGCGCCCGGGGGCGGCAACAGGAGCTGAGGGTCACTTTGTGGCAGG 623
573 TGCCGTGGAGCGCCCAAGAGCGCCCTGCGCTAGAGCGCTGCTCTGGAGGGCTGCGCGCG 632
624 TGCCGTGGAGCGCCCAAGAGCGCCCTGCGCTAGAGCGCTGCTCTGGAGGGCTGCGCGCG 683
633 CTTCTCATTTGCGAGCTTCAAGGAGCTGAGAGCGCGCTGGAGCAACCCCGTGAACGGGCA 692
684 CTTCTCATTTGCGAGCTTCAAGGAGCTGAGAGCGCGCTGGAGCAACCCCGTGAACGGGCA 743
693 GCCAGCCAGTGGATGACCTTCTCATCTCTACTGGGGTGGAGAGATCGGACAGAGAT 752
744 GCCAGCCAGTGGATGACCTTCTCATCTCTACTGGGGTGGAGAGATCGGACAGAGAT 803
753 CGCAGAGTCAAGGAGCTGCTTCCACTGCCACGCTTTCCTCCGTTTCTGACAGAGAGGC 812
804 CGCAGAGTCAAGGAGCTGCTTCCACTGCCACGCTTTCCTCCGTTTCTGACAGAGAGGC 863
813 CGGCTTGGGGCTTGGAGAGCTGCAACAGAGAGCGAGAGCTGCAAGAGGTCTCTCG 872
864 CGGCTTGGGGCTTGGAGAGCTGCAACAGAGAGCGAGAGCTGCAAGAGGTCTCTCG 923
873 GGAGACAGCGGTCTTGGAGCGAGTGTAGCGCGGTGCTGCAAGCTGCTGCGCGCAGG 932
924 GGAGACAGCGGTCTTGGAGCGAGTGTAGCGCGGTGCTGCAAGCTGCTGCGCGCAGG 983
933 CGAGGTGCAAGTCCACAGATGAAGCGCTGTACCTGGCCCTGAAACCAAGTGCAGCGTGAG 992
984 CGAGGTGCAAGTCCACAGATGAAGCGCTGTACCTGGCCCTGAAACCAAGTGCAGCGTGAG 1043
993 CACACGCAAGTGGCTCATTTGCGAGCGCTGGTGTCTGTGGAGAGCTGCGCGCGCT 1052
1044 CACACGCAAGTGGCTCATTTGCGAGCGCTGGTGTCTGTGGAGAGCTGCGCGCGCT 1103
1053 GCAGAGGCGCTGCGGAGCAGCTCGATGGAGGAGGAGTGAAGTGGGCTCAGCGCAT 1112
1104 GCAGAGGCGCTGCGGAGCAGCTCGATGGAGGAGGAGTGAAGTGGGCTCAGCGCAT 1163
1113 CCGCTGCGGAGCATGCGCCCGCCACACTCATTCGCGACCAACCGCTTCAAGGCGAGCTTCCA 1172
1164 CCGCTGCGGAGCATGCGCCCGCCACACTCATTCGCGACCAACCGCTTCAAGGCGAGCTTCCA 1223
1173 GGGCATGTGGATCGGTACGGCGTGGCGCTTACAGAGGTCATCCCGCTCCCTACAC 1232
1224 GGGCATGTGGATCGGTACGGCGTGGCGCTTACAGAGGTCATCCCGCTCCCTACAC 1283
1233 CATCATACCTTCCCTTCTGTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1292
1284 CATCATACCTTCCCTTCTGTTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1343
1293 CATGTTCTCTTTCGCGCTGGCCATGGTCTCTTGGAGAACCGAGCGGCTGTGAAAGCGCG 1352

1344 CATGTTCTCTTTCGCGCTGGCCATGGTCTCTGCGAGAACCGAGCGGCTGTGAAGCGCGC 1403
1353 GCAGAACGAGATCTGGCAGACTTTCTTCAGGGGCGCGCTACCTGCTCTCTTATATGGGCT 1412
1404 GCAGAACGAGATCTGGCAGACTTTCTTCAGGGGCGCGCTACCTGCTCTCTTATGGGCT 1463
1413 GTTCTCATCTACACGCGCTTCACTACACGAGTGTCTTCACTGCGGCGCACGAGCATCTT 1472
1464 GTTCTCATCTACACGCGCTTCACTACACGAGTGTCTTCACTGCGGCGCACGAGCATCTT 1523
1473 CCGCTTGGGCTGAGTGTGGCGCGCATGGCCAAACAGTCTGGCTGGAGTGTGATCTCT 1532
1524 CCGCTTGGGCTGAGTGTGGCGCGCATGGCCAAACAGTCTGGCTGGAGTGTGATCTCT 1583
1533 GGCAGACACAGATGTCTTACCTCGATGCCAAACGTCACCGGTGTCTTCTGGGACCCCTA 1592
1584 GGCAGACACAGATGTCTTACCTCGATGCCAAACGTCACCGGTGTCTTCTGGGACCCCTA 1643
1593 CCGCTTGGCAGTGTCTTATTTGGAGCTGCTGCGCAACCACTTGGAGCTTCTCACTC 1652
1644 CCGCTTGGCAGTGTCTTATTTGGAGCTGCTGCGCAACCACTTGGAGCTTCTCACTC 1703
1653 CTTCAAGATGAAGATGTCCGTCTTCTGGGCGCTGCTGCAATGGCTTTGGGGTGTCTCT 1712
1704 CTTCAAGATGAAGATGTCCGTCTTCTGGGCGCTGCTGCAATGGCTTTGGGGTGTCTCT 1763
1713 CGGAGTCTTCAACCAAGTGTCTTGGGCGAGAGCAACCGGTGTCTGCTGGAGACGCTGCC 1772
1764 CGGAGTCTTCAACCAAGTGTCTTGGGCGAGAGCAACCGGTGTCTGCTGGAGACGCTGCC 1823
1773 GGAGTCACTTCTGCTGGGACTCTTGGGTACTCTGCTGTCTTCTAGTCACTACAGTG 1832
1824 GGAGTCACTTCTGCTGGGACTCTTGGGTACTCTGCTGTCTTCTAGTCACTACAGTG 1883
1833 GCTGTGTCTTGGGCTGCGAGGGCGCGCTCG ---CCAGACTCTCTCATCTTCACTCAA 1889
1884 GCTGTGTCTTGGGCTGCGAGGGCGCGCTCGGGCGCGAGCTCTCATCTCACTCAA 1943
1890 CATGTTCTTCTTCTCCCAAGCGCCAGCAACAGGTGTCTTACCCCGGAGAGGTGT 1949
1944 CATGTTCTTCTTCTCCCAAGCGCCAGCAACAGGTGTCTTACCCCGGAGAGGTGT 2003
1950 CCAGGCGAGCTGTGTGCTGCGCTTGGCCATGTGTGCGCATCTGCTGCTTGGCACACC 2009
2004 CCAGGCGAGCTGTGTGCTGCGCTTGGCCATGTGTGCGCATCTGCTGCTTGGCACACC 2063
2010 CCGTCACTTGTGCAACCGCCAGCGCGCTGCGAGAGGGCGCGCTGACCGAGCA 2069
2064 CCGTCACTTGTGCAACCGCCAGCGCGCTGCGAGAGGGCGCGCTGACCGAGCA 2123
2070 GGAAACAAAGCGCGGCTTGTGGACCTGCTGACCGCATCTGTGAATGGCTGGAGTCCGA 2129
2124 GGAAACAAAGCGCGGCTTGTGGACCTGCTGACCGCATCTGTGAATGGCTGGAGTCCGA 2183
2130 TGAGAAAGGCGAGGGCGCTGGATGAAGAGAGGGCGAGCTGCTCCCTCCGAGGT 2189
2184 TGAGAAAGGCGAGGGCGCTGGATGAAGAGAGGGCGAGCTGCTCCCTCCGAGGT 2243
2190 GCTCATGCAACGAGCGCATCCACCATCGAGTTCTTGGCTGGGCTGCTCTCCAAACCGC 2249
2244 GCTCATGCAACGAGCGCATCCACCATCGAGTTCTTGGCTGGGCTGCTCTCCAAACCGC 2303
2250 TCTCTACCTGGCGCTGTGGGCGCTGAGCTTGGCCCAACCGCCAGCTGTGCGAGGTTCTG 2309
2304 CTCTTACCTGGCGCTGTGGGCGCTGAGCTTGGCCCAACCGCCAGCTGTGCGAGGTTCTG 2363
2310 GGCATGTGTGATGCGCATAGGCTTGGGCTGGGCGGAGGTGGGCTGGGCTGGGCTGGG 2369
2364 GGCATGTGTGATGCGCATAGGCTTGGGCTGGGCGGAGGTGGGCTGGGCTGGGCTGGG 2423
2370 GCTGTGCTCCCATCTTGTGCGCGCTTGTGCGGTGATGACCGGTGGCTATCTCTGCTGGTGA 2429

Db 2424 GCTGGTCCCATCTTTGGCGCCTTTGGCGTGTGATGACCGTGGCTATCTCTGGTGTGATGGA 2483
Qy 2430 GGGACTCTCAGCCTTCTGCGACGCGCTGCGGCTGCACTGGGTGAAATCCAGAAACAAGTT 2489
Db 2484 GGGACTCTCAGCCTTCTGCGACGCGCTGCGGCTGCACTGGGTGAAATCCAGAAACAAGTT 2543
Qy 2490 CTACTCAGGACGCGGCTACAGCTGAGTCCCTTCCACCTTCCGCTGCCACAGATGACTAGG 2549
Db 2544 CTACTCAGGACGCGGCTACAGCTGAGTCCCTTCCACCTTCCGCTGCCACAGATGACTAGG 2603
Qy 2550 CCCACTCAGTCTGCGCAGACCTCTTCTGACCTCTGAGGAGGAGGAGGAAATAAGAC 2609
Db 2604 CCCACTCAGTCTGCGCAGACCTCTTCTGACCTCTGAGGAGGAGGAGGAAATAAGAC 2663
Qy 2610 GGTCCGCGCTGGCAAAAAA 2640
Db 2664 GGTCCGCGCTGGCAGTGAAAAA 2694

RESULT 2

US-11-266-748A-76629
; Sequence 76629, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 5815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76629
; LENGTH: 1696
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-76629

Query Match 55.8%; Score 1474.2; DB 7; Length 1696;

Best Local Similarity 99.1%; Pred. No. 1.3e-255;

Matches 1504; Conservative 0; Mismatches 8; Indels 5; Gaps 2;

Qy 781 CAGTCTTCCCGTTCTGACGACGAGGAGGCGCGCTCGGGCCCTGACGAGCTGCAA 840
Db 28 CAGTCTTCCCGTTCTGACGACGAGGAGGCGCGCTCGGGCCCTGACGAGCTGCAA 87
Qy 841 CACGACGACGAGCTGACGAGGAGTCTCGGGAGACAGAGCGTTCTTGAGCAGGTG 900
Db 88 CACGACGACGAGCTGACGAGGAGTCTCGGGAGACAGAGCGTTCTTGAGCAGGTG 147
Qy 901 CTAGGCGGGTGTGAGCTGCTGCGCCAGGAGGAGTGCAGATGAAAGGCC 960
Db 148 CTAGGCGGGTGTGAGCTGCTGCGCCAGGAGGAGTGCAGATGAAAGGCC 207
Qy 961 GTGTACTTGGCCCTTGACCAAGTGCAGGTGAGCACACGCAAGTGCCTCATTTGCCGAG 1020

Db 208 GTGTACTGGCCCTGAAACCAAGTGCAGCGTGAGCAACCAAGTGCCTCATTTGCCGAG 267
Qy 1021 GCCTGGTGTCTGTGCGAGACCTGCCCGCCTGCGAGGCGCTGCGGAGCAGTCCGATG 1080
Db 268 GCCTGGTGTCTGTGCGAGACCTGCCCGCCTGCGAGGCGCTGCGGAGCAGTCCGATG 327
Qy 1081 GAGAGGAGTGTGAGTGTGCGGTCAACCGATCCCTGCGGGGACATGCCCCCACAATC 1140
Db 328 GAGAGGAGTGTGAGTGTGCGGTCAACCGATCCCTGCGGGGACATGCCCCCACAATC 387
Qy 1141 ATCCGACCAACCGCTTTCAGCGGCGCTTCCAGGCGCATCTGAGTCCCTGCGGCGTGGC 1200
Db 388 ATCCGACCAACCGCTTTCAGCGGCGCTTCCAGGCGCATCTGAGTCCCTGCGGCGTGGC 447
Qy 1201 CGCTTACGAGGAGTCAACCGCGCTTCCATACCATCATCTTCCCTTCTCTTTTGTCT 1260
Db 448 CGCTTACGAGGAGTCAACCGCGCTTCCATACCATCATCTTCCCTTCTCTTTTGTCT 507
Qy 1261 GTGATGTTCGGGATGTGGGCGACGCGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1320
Db 508 GTGATGTTCGGGATGTGGGCGACGCGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 567
Qy 1321 CTTGCGGAGAACCGACCGGCTGTGAAGCGCGCAGAACGAGATCTGGCAGACTTCTTCT 1380
Db 568 CTTGCGGAGAACCGACCGGCTGTGAAGCGCGCAGAACGAGATCTGGCAGACTTCTTCT 627
Qy 1381 AGGGCGCGTACCTGTCTTCTTATGGGCGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440
Db 628 AGGGCGCGTACCTGTCTTCTTATGGGCGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCT 687
Qy 1441 AAGAGTGTCTTCTGCGGCGACAGCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1500
Db 688 AAGAGTGTCTTCTGCGGCGACAGCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 747
Qy 1501 GCCAACCAAGTCTGGCTGGAGTGTGATTCCTTGGGCGCGACAGCATCTTCTTCTTCTTCT 1560
Db 748 GCCAACCAAGTCTGGCTGGAGTGTGATTCCTTGGGCGCGACAGCATCTTCTTCTTCTTCT 807
Qy 1561 CCCAACGTCACCGGTGTCTTCTTGGGCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1620
Db 808 CCCAACGTCACCGGTGTCTTCTTGGGCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 867
Qy 1621 CTGGCTGCCAACCACTTGTAGCTTCTCAACTCTTCTTCAAGATGAAGATGTCGTCATCTG 1680
Db 868 CTGGCTGCCAACCACTTGTAGCTTCTCAACTCTTCTTCAAGATGAAGATGTCGTCATCTG 927
Qy 1681 GCGGTGTGCACATGCGCTTTTGGGTGTGTGCTTCTTCAAGATGAAGATGTCGTCATCTG 1738
Db 928 GCGGTGTGCACATGCGCTTTTGGGTGTGTGCTTCTTCAAGATGAAGATGTCGTCATCTG 987
Qy 1739 GCCAGAGGCAACCGGCTGTCTGAGACGCTGCGGAGCTTCTTCTTCTTCTTCTTCTTCTTCT 1798
Db 988 GCCAGAGGCAACCGGCTGTCTGAGACGCTGCGGAGCTTCTTCTTCTTCTTCTTCTTCTTCT 1047
Qy 1799 TCGGTACTCTGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1858
Db 1048 TCGGTACTCTGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1107
Qy 1859 CTTCTG---CCGACATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1915
Db 1108 CTTCTGCCCCCAGCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1167
Qy 1916 GCAACAGGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1975
Db 1168 GCAACAGGCTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1227
Qy 1976 TGCCCATGTGTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2035
Db 1228 TGCCCATGTGTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1287
Qy 2036 GCGGCTTGGGAGGCGCGCTGACGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2095

Db 1288 GCCGCTGCGGAGGAGCCGCTGACCGACAGAGGAAACAAGGCCGGGTGCTGGACC 1347
Qy 2096 TGCTGACGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGGAGGGGGCTGGATG 2155
Db 1348 TGCTGACGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGGAGGGGGCTGGATG 1407
Qy 2156 ATGAGAGGAGGCGGAGCTGCTCCCTCCGAGGTGCTCATGCCAGGCGCATCCACACCA 2215
Db 1408 ATGAGAGGAGGCGGAGCTGCTCCCTCCGAGGTGCTCATGCCAGGCGCATCCACACCA 1467
Qy 2216 TCGAGTTCTGCTGGGCTGGCTTCCAAACCCGCTCTTACCTGCGGCTGTGGGCCCTGA 2275
Db 1468 TCGAGTTCTGCTGGGCTGGCTTCCAAACCCGCTCTTACCTGCGGCTGTGGGCCCTGA 1527
Qy 2276 GCCTGGCCACGCGCCAG 2292
Db 1528 GCCTGGCCACGCGCCG 1544

RESULT 4

US-11-266-748A-129440/c
; Sequence 129440, Application US/11266748A
; Publication No. US20060134663A1

GENERAL INFORMATION:

; APPLICANT: Harkin, Paul
; APPLICANT: Mulligan, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A

; CURRENT FILING DATE: 2005-11-03

; PRIOR APPLICATION NUMBER: EP 04105479.2

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105482.6

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105483.4

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105507.0

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105485.9

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105484.2

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: US 60/662,276

; PRIOR FILING DATE: 2005-03-14

; PRIOR APPLICATION NUMBER: US 60/700,293

; PRIOR FILING DATE: 2005-07-18

; NUMBER OF SEQ ID NOS: 483996

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 129440

; LENGTH: 1696

; TYPE: DNA

; ORGANISM: Homo Sapiens

US-11-266-748A-129440

Query Match 55.8%; Score 1474.2; DB 7; Length 1696;

Best Local Similarity 99.1%; Pred. No. 1.3e-255;

Matches 1504; Conservative 0; Mismatches 8; Indels 5; Gaps 2;

Qy 781 CAGCTTCCCGTTTCTGACGAGGAGGAGGCCGCGCTCGGGGCGCTGACGAGCTGCAA 840
Db 1669 CAGCTTCCCGTTTCTGACGAGGAGGAGGCCGCGCTCGGGGCGCTGACGAGCTGCAA 1610
Qy 841 CAGCAGAGCAGGAGCTGACGAGGTCTCTCGGGAGACAGAGCGGTCTTCTGAGCCAGGTG 900
Db 1609 CAGCAGAGCAGGAGCTGACGAGGTCTCTCGGGAGACAGAGCGGTCTTCTGAGCCAGGTG 1550
Qy 901 CTAGGCGGGGTGCTGAGCTGTGCTGCCGCCAGGCGAGGTGCAAGATGAAGGCC 960
Db 1549 CTAGGCGGGGTGCTGAGCTGTGCTGCCGCCAGGCGAGGTGCAAGATGAAGGCC 1490
Qy 961 GTGTACTGCGCTTGAAACGAGTGCAGGTGAGCACCAGCAAGTGCCTCATTTGCCGAG 1020

Db 1489 GTGTACTGCGCTGAAACCAAGTGCAGCTGAGCACCAAGTGCCTCATTTGCCGAG 1430
Qy 1021 GCCTGGTGTCTGTGCGAGACCTGCCCGCTTGAGGAGCCCTGCGGAGCAGCTCGATG 1080
Db 1429 GCCTGGTGTCTGTGCGAGACCTGCCCGCTTGAGGAGCCCTGCGGAGCAGCTCGATG 1370
Qy 1081 GAGGAGGAGTGTAGTGCCTGTCACCGCATCCCTGCGGGGACATGCCCCCACAATC 1140
Db 1369 GAGGAGGAGTGTAGTGCCTGTCACCGCATCCCTGCGGGGACATGCCCCCACAATC 1310
Qy 1141 ATCGCACCAACCGCTTCAAGGCGAGCTTCCAGGCGATCGTGTGCTGAGGCGTGGC 1200
Db 1309 ATCGCACCAACCGCTTCAAGGCGAGCTTCCAGGCGATCGTGTGATGCTTACGCGGTGGC 1250
Qy 1201 CGCTACCGAGAGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCTGTTGCT 1260
Db 1249 CGCTACCGAGAGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCTGTTGCT 1190
Qy 1261 GTGATGTTGGGAGTGTGGCCACGCGGTCTCATGTTCTTCTGCGCCCTGGCCATGCTC 1320
Db 1189 GTGATGTTGGGAGTGTGGCCACGCGGTCTCATGTTCTTCTGCGCCCTGGCCATGCTC 1130
Qy 1321 CTTGCGGAGAACCGACCGGCTGTAAAGCCGCGACAGAACGAGATCTGGCAGACTTCTTC 1380
Db 1129 CTTGCGGAGAACCGACCGGCTGTAAAGCCGCGACAGAACGAGATCTGGCAGACTTCTTC 1070
Qy 1381 AGGGGCGCTTACCTGCTCTGCTTATGGGCTGTCTTCAATCTACACCGGCTTCACTAC 1440
Db 1069 AGGGGCGCTTACCTGCTCTGCTTATGGGCTGTCTTCAATCTACACCGGCTTCACTAC 1010
Qy 1441 AACGAGTGTTCAGTCCGCGCACAGCATCTTCCCTCGGGCTGGAGTGTGGCGGCATG 1500
Db 1009 AACGAGTGTTCAGTCCGCGCACAGCATCTTCCCTCGGGCTGGAGTGTGGCGGCATG 950
Qy 1501 GCCAACGAGTCTGGCTGGAGTGTATGATTCCTGGCCGACACAGATGCTTACCTGGAT 1560
Db 949 GCCAACGAGTCTGGCTGGAGTGTATGATTCCTGGCCGACACAGATGCTTACCTGGAT 890
Qy 1561 CCCAAGCTCACCGGTGTCTTCTGGGACCTACCCCTTTGGCATCGATCTTATTTGAGC 1620
Db 889 CCCAAGCTCACCGGTGTCTTCTGGGACCTACCCCTTTGGCATCGATCTTATTTGAGC 830
Qy 1621 CTGGCTGCCAACCACTTGAGCTTCTCAACTCTTCAAGATGAAGATGCTCGTCACTCTG 1680
Db 829 CTGGCTGCCAACCACTTGAGCTTCTCAACTCTTCAAGATGAAGATGCTCGTCACTCTG 770
Qy 1681 GCGCTGTCACATGGCCTTT--GGGGTGGTCTCGAGTCTTCAACCAAGTGCACCTTG 1738
Db 769 GCGCTGTCACATGGCCTTTTGGGGTGGGTCTCGAGTCTTCAACCAAGTGCACCTTG 710
Qy 1739 GCCAGAGGACCGGCTGCTGGAGACGCTGCGGAGCTCACCTTCTGCTGGGACTCT 1798
Db 709 GCCAGAGGACCGGCTGCTGGAGACGCTGCGGAGCTCACCTTCTGCTGGGACTCT 650
Qy 1799 TGGTTTACTCGGTTCCTAGTCACTCAAGTGGTGTGTGTGGGTGCGAGGCGC 1858
Db 649 TGGTTTACTCGGTTCCTAGTCACTCAAGTGGTGTGTGTGGGTGCGAGGCGC 590
Qy 1859 CTTGCGCCCGGAGCATCTTCACTTCAACATGTTCTTCTTCTCCACAGCCCCA 1915
Db 589 CTTGCGCCCGGAGCATCTTCACTTCAACATGTTCTTCTTCTCCACAGCCCCA 530
Qy 1916 GCAACAGGTGTCTTACCCCGGAGGAGTGTCCAGGCGAGCTTCTTCTCCACAGCCCCA 1975
Db 529 GCAACAGGTGTCTTACCCCGGAGGAGTGTCCAGGCGAGCTTCTTCTTCTTCTGCGCT 470
Qy 1976 TGGCCATGTGCGCATCTGCTGTTGGCAACCCCTGCACTGCTGCAACCGCCACCGCC 2035
Db 469 TGGCCATGTGCGCATCTGCTGTTGGCAACCCCTGCACTGCTGCAACCGCCACCGCC 410
Qy 2036 GCGCCCTGCGGAGGAGCGGCTGACGAGAGGAAACAAGGCGGGTTCCTGGACC 2095

Db 409 GCCGCTCGGAGGAGGCCGCTGACCGACAGAGGAAAAAAGGCGCGGTTGCTGGACC 350
QY 2096 TGCCTGACCATCTGTAATGCTGAGCTCCGATGAGGAAAGGAGGGGCGCTGGATG 2155
Db 349 TGCCTGACCATCTGTAATGCTGAGCTCCGATGAGGAAAGGAGGGGCGCTGGATG 290
QY 2156 ATGAAGAGGAGGCGGAGCTCGTCCCTCCGAGGTGCTCATGACACAGGCGCATCCACCA 2215
Db 289 ATGAAGAGGAGGCGGAGCTCGTCCCTCCGAGGTGCTCATGACACAGGCGCATCCACCA 230
QY 2216 TCGAGTTCTGCTGGGCTGGGCTCGGCTCCAAACACCGCTCCCTACCTGCGGCTGTGGGCGCTGA 2275
Db 229 TCGAGTTCTGCTGGGCTGGGCTCGGCTCCAAACACCGCTCCCTACCTGCGGCTGTGGGCGCTGA 170
QY 2276 GCCTGGCCCAAGCCGAG 2292
Db 169 GCCTGGCCCAAGCCGAG 153

RESULT 5

US-11-266-748A-361154
; Sequence 361154, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIORITY FILING DATE: 2005-11-03
; PRIORITY APPLICATION NUMBER: EP 04105479.2
; PRIORITY FILING DATE: 2004-11-03
; PRIORITY APPLICATION NUMBER: EP 04105482.6
; PRIORITY FILING DATE: 2004-11-03
; PRIORITY APPLICATION NUMBER: EP 04105483.4
; PRIORITY FILING DATE: 2004-11-03
; PRIORITY APPLICATION NUMBER: EP 04105507.0
; PRIORITY FILING DATE: 2004-11-03
; PRIORITY APPLICATION NUMBER: EP 04105485.9
; PRIORITY FILING DATE: 2004-11-03
; PRIORITY APPLICATION NUMBER: EP 04105484.2
; PRIORITY FILING DATE: 2004-11-03
; PRIORITY APPLICATION NUMBER: US 60/662,276
; PRIORITY FILING DATE: 2005-03-14
; PRIORITY APPLICATION NUMBER: US 60/700,293
; PRIORITY FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 361154
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1709)..(1709)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1799)..(1799)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1803)..(1803)
; OTHER INFORMATION: n is a, c, g, or t

US-11-266-748A-361154

Query Match 41.4%; Score 1093.2; DB 7; Length 1839;
Best Local Similarity 98.6%; Pred. No. 2.9e-187;
Matches 1157; Conservative 8; Mismatches 9; Gaps 5;
QY 781 CACGTCTTCCCGTTTCTGACGAGGAGGCGCCGCTCGGGGCGCTGACGAGCTGCAA 840

Db 28 CACGTCTTCCCGTTTCTGACGAGGAGGCGCCGCTCGGGGCGCTGACGAGCTGCAA 87
QY 841 CAGCAGAGCCAGAGCTGACAGAGTCTCTCGGGGAGACAGAGCGGTCTCTGAGCCAGGTG 900
Db 88 CAGCAGAGCCAGAGCTGACAGAGTCTCTCGGGGAGACAGAGCGGTCTCTGAGCCAGGTG 147
QY 901 CTAGGCGGGGTCTGACAGTCTGCGGCCAGGCGGAGGTGACAGTCCACAAGATGAAGGCC 960
Db 148 CTAGGCGGGGTCTGACAGTCTGCGGCCAGGCGGAGGTGACAGTCCACAAGATGAAGGCC 207
QY 961 GTGTACCTTGGCCCTGAACCAAGTGCAGCGTGCAGCACAAGTGCCTCATTTGCCGAG 1020
Db 208 GTGTACCTTGGCCCTGAACCAAGTGCAGCGTGCAGCACAAGTGCCTCATTTGCCGAG 267
QY 1021 GCCTGGTCTCTGTGCGAGACCTGCCCGCTGCGAGGAGCGCTGCGGACAGTCCGATG 1080
Db 268 GCCTGGTCTCTGTGCGAGACCTGCCCGCTGCGAGGAGCGCTGCGGAGACAGTCCGATG 327
QY 1081 GAGGAGGAGTGAAGTCCGCTGAGTCAACCGCATCCCTGCGGGAGACATGCCCCCACATC 1140
Db 328 GAGGAGGAGTGAAGTCCGCTGAGTCAACCGCATCCCTGCGGGAGACATGCCCCCACATC 387
QY 1141 ATCCCGCACCAACCGCTTCAAGCGCAGCTTCCAGGGCATCGTGGATCGCTACGCGTGGGC 1200
Db 388 ATCCCGCACCAACCGCTTCAAGCGCAGCTTCCAGGGCATCGTGGATCGCTACGCGTGGGC 447
QY 1201 CGCTACGAGGAGTCAACCGCGCTCCCTACACCATCATCATCCCTTCCCTTCTGTTGCT 1260
Db 448 CGCTACGAGGAGTCAACCGCGCTCCCTACACCATCATCATCCCTTCCCTTCTGTTGCT 507
QY 1261 GTGATGTTTC-GGGGATGTGGGCGACAGGCTGCTCATGTCTCTTTCGCGCTGGCATGTG 1319
Db 508 GTGATGTTTCGGGGGATGTGGGCGACGCGCTGCTCATGTCTCTTTCGCGCTGGCATGTG 567
QY 1320 CCTTGGGAGAACCGACCGGCTGTGAAGCGCGGAGACGAGATCTGGCAGACTTCTT 1379
Db 568 CCTTGGGAGAACCGACCGGCTGTGAAGCGCGGAGACGAGATCTGGCAGACTTCTT 627
QY 1380 CAGGGGCGCTACCTGCTGCTGCTTATGGGCGCTGTTTCCATCTACACGGGTTCACTA 1439
Db 628 CAGGGGCGCTACCTGCTGCTGCTTATGGGCGCTGTTTCCATCTACACGGGTTCACTA 687
QY 1440 CAAAGAGTGTTCAGTTCGGCGCACAGCATCTTCCCGCTGGGCTGGAGTGTGGCGGC 1497
Db 688 CAAAGAGTGTTCAGTTCGGCGCACAGCATCTTCCCGCTGGGCTGGAGTGTGGCGGC 747
QY 1498 ATGGCCAAACAGTCTGGCTGGAGTGAATGCAATTCCTTGGCGGAGCACACGATGCTTACCCCTG 1557
Db 748 ATGGCCAAACAGTCTGGCTGGAGTGAATGCAATTCCTTGGCGGAGCACACGATGCTTACCCCTG 807
QY 1558 GATCCAAAGTCAACGGTGTCTTCTGGGACCCCTACCCCTTTGGCATCGATCTATTTGG 1617
Db 808 GATCCAAAGTCAACGGTGTCTTCTGGGACCCCTACCCCTTTGGCATCGATCTATTTGG 867
QY 1618 AGCTGGGCTGCCAACCACTTGAAGTCTCTCAACTCTTCAAGATGAAGATGTCGCTCATC 1677
Db 868 AGCTGGGCTGCCAACCACTTGAAGTCTCTCAACTCTTCAAGATGAAGATGTCGCTCATC 927
QY 1678 CTGGGCGTGTGCAATGGCCCTTT--GGGGTGGTCTCTGGAGTCTTCAACCAAGTGCATC 1735
Db 928 CTGGGCGTGTGCAATGGCCCTTTTGGGGTGGTCTCTGGAGTCTTCAACCAAGTGCATC 987
QY 1736 TTGGCCAGAGGACCGGCTGCTGAGAGACGCTCGGAGTCACTTCTCTGCTGGGAC 1795
Db 988 TTGGCCAGAGGACCGGCTGCTGAGAGACGCTCGGAGTCACTTCTCTGCTGGGAC 1047
QY 1796 TC-TTCGGTTACCTCGTGTCTAGTCACTACAAAGTGGCTGTGTGTGCTGGGCTGCCAGG 1854
Db 1048 TCTTTTCGGTTACCTCGTGTCTAGTCACTACAAAGTGGCTGTGTGTGCTGGGCTGCCAGG 1107
QY 1855 GCGCGCTCG---CCAGCATCTCATCACTTCAATCAACATGTTCTCTTCTTCTCCACAGC 1911

Db 1108 GCGCCTCGGCCCCCAGCATCTTCATCCACTTCAACATGTTCTCTTCTCCACAGC 1167
Qy 1912 CCAGCAACAGGCTGCTTACCCCGCAGGAGG 1945
Db 1168 CCAGCAACAGGCTGCTTACCCCGCAGGAGG 1201

RESULT 6

US-11-266-748A-444533/c
; Sequence 444533, Application US/11266748A
; Publication No. US20060134663A1

; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A

; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 444533
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (37)..(37)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (41)..(41)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131)..(131)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-444533

Query Match 41.4%; Score 1093.2; DB 7; Length 1839;
Best Local Similarity 98.6%; Pred. No. 2.9e-187;
Matches 1157; Conservative 0; Mismatches 8; Indels 9; Gaps 5;

Qy 781 CAGTCTTCCGTTCTGCAGCAGGAGGAGGCGCGCTCGGGCGCTGCAGCAGCTGCAA 840
Db 1812 CAGTCTTCCGTTCTGCAGCAGGAGGAGGCGCGCTCGGGCGCTGCAGCAGCTGCAA 1753
Qy 841 CAGCAGAGCCAGGAGCTGCAGGAGGCTCTCGGGGAGACAGAGCGGTTCTTGAGCCAGGTG 900
Db 1752 CAGCAGAGCCAGGAGCTGCAGGAGGCTCTCGGGGAGACAGAGCGGTTCTTGAGCCAGGTG 1693
Qy 901 CTAGGCGGGGTGCTGCAGTGTGCGGCCAGGCGAGGTGCCACAGATGAAGGCC 960
Db 1692 CTAGGCGGGGTGCTGCAGTGTGCGGCCAGGCGAGGTGCCACAGATGAAGGCC 1633
Qy 961 GTGTACTCTGGCCTGTAACCAAGTCAGCGGTGAGCACCAGCAAGTGCCTCATTTGCCGAG 1020

RESULT 7

US-11-266-748A-221798
; Sequence 221798, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:

Db 1632 GTGTACCTGGCCCTCGAAACCAAGTGCAGCACCAGCAAGTGCCTCATTTGCCAG 1573
Qy 1021 GCCTGGTGTCTGTGCGAGACCTGCCCGCTGCAGGAGGCGCTCGGGACAGTCCGATG 1080
Db 1572 GCCTGGTGTCTGTGCGAGACCTGCCCGCTGCAGGAGGCGCTCGGGACAGTCCGATG 1513
Qy 1081 GAGGAGGAGTGTGCTGCGGTGTCAACCGATCCCTCGCCGGGACATGCCCCCACAATC 1140
Db 1512 GAGGAGGAGTGTGCTGCGGTGTCAACCGATCCCTCGCCGGGACATGCCCCCACAATC 1453
Qy 1141 ATCGCACCAACCGCTTTCAGGCGAGCTTCCAGGGGATCGTGGATCGTACGGGCTGGGC 1200
Db 1452 ATCGCACCAACCGCTTTCAGGCGAGCTTTCAGGGGATCGTGGATCGTACGGGCTGGGC 1393
Qy 1201 CGCTTACCGAGAGTCAACCCCGCTTCCATACCATCATCATCTTCCCTTCCCTCTTGTGCT 1260
Db 1392 CGCTTACCGAGAGTCAACCCCGCTTCCATACCATCATCATCTTCCCTTCCCTCTTGTGCT 1333
Qy 1261 GTGATGTTTC-GGGGATGTGGGCAACCGGCTGTGATGTTCTTCCCTCGCCCTGGCCATGCT 1319
Db 1332 GTGATGTTTCGGGGGATGTGGGCAACCGGCTGTGATGTTCTTCCCTCGCCCTGGCCATGCT 1273
Qy 1320 CCTTGGCGGAAACCGACCGGCTGTGAAGCGCGGCAAGAGATCTGGCAGACTTCTT 1379
Db 1272 CCTTGGCGGAAACCGACCGGCTGTGAAGCGCGGCAAGAGATCTGGCAGACTTCTT 1213
Qy 1380 CAGGGGCGGCTTACCTGCTCTGTGTTATGGGCTGTGTTCTCCATCTACACCGGCTTCACTA 1439
Db 1212 CAGGGGCGGCTTACCTGCTCTGTGTTATGGGCTGTGTTCTCCATCTACACCGGCTTCACTA 1153
Qy 1440 CAACGAGTGTCTCAGTGGGCGCAACAGATCTTCCCG--TCGGGCTGGAGTGGCGGCC 1497
Db 1152 CAACGAGTGTCTCAGTGGGCGCAACAGATCTTCCCGCTCGGGCTGGAGTGGCGGCC 1093
Qy 1498 ATGGCCAAACAGTCTGGCTGGAGTGTGATTCCTGGGCCAGCAGATGCTTACCCTG 1557
Db 1092 ATGGCCAAACAGTCTGGCTGGAGTGTGATTCCTGGGCCAGCAGATGCTTACCCTG 1033
Qy 1558 GATCCCAACAGTCAACCGGCTGTCTTCTGGGAGCCCTACCCCTTTCGATCGATCCTATTG 1617
Db 1032 GATCCCAACAGTCAACCGGCTGTCTTCTGGGAGCCCTACCCCTTTCGATCGATCCTATTG 973
Qy 1618 AGCTTGGCTGCCAAACCACTTGAAGTCTCTCAACTCTTCAAGATGAAGATGTCCGTCATC 1677
Db 972 AGCTTGGCTGCCAAACCACTTGAAGTCTCTCAACTCTTCAAGATGAAGATGTCCGTCATC 913
Qy 1678 CTGGGCGTGTGCACATGGCTTTT--GGGCTGGTCTCTCGAGTCTTCAACCAAGTGCAC 1735
Db 912 CTGGGCGTGTGCACATGGCTTTTGGGGTGGGCTCTCGAGTCTTCAACCAAGTGCAC 853
Qy 1736 TTGGCCAGAGGCAACCGGCTGTCTGAGAGCGCTGCGGAGCTCACCTTCTCTGCTGGAC 1795
Db 852 TTGGCCAGAGGCAACCGGCTGTCTGAGAGCGCTGCGGAGCTCACCTTCTCTGCTGGAC 793
Qy 1796 TC-TTCGTTTACCTGTGTTCTTGTGATCATCTACAAGTGGCTGTGTGTCTGGGCTGCCAG 1854
Db 792 TCTTTTCGTTTACCTGTGTTCTTGTGATCATCTACAAGTGGCTGTGTGTCTGGGCTGCCAG 733
Qy 1855 GCGGCTCTG---CCAGCATCTCTCATCTCCACTTTCATCAACATGTTCTCTTCTCCACAGC 1911
Db 732 GCGGCTCTGCGGCTCCAGCATCTCTCATCTCCACTTTCATCAACATGTTCTCTTCTCCACAGC 673
Qy 1912 CCAGCAACAGGCTGTCTTACCCCGCGGAGGAGG 1945
Db 672 CCAGCAACAGGCTGTCTTACCCCGCGGAGGAGG 639

```
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 221798
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-221798
```

Query Match 36.9%; Score 973.2; DB 7; Length 1000;
Best Local Similarity 99.4%; Pred. No. 8.5e-166;
Matches 988; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

| QY | Db | 541 | 2247 | 601 | 2307 | 661 | 2367 | 721 | 2427 | 781 | 2487 | 841 | 2547 | 901 | 2607 | 961 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----|----|--|------|-----|--|-----|------|--|------|------|--|-----|------|--|------|-----|--|----|-----|--|----|------|--|----|------|--|----|-----|--|----|-----|--|----|------|--|----|------|--|----|-----|--|----|-----|--|----|------|--|----|------|--|----|-----|--|----|-----|--|----|------|--|----|------|--|----|-----|--|----|-----|--|----|------|--|----|------|--|----|-----|--|----|-----|--|----|------|--|----|------|--|----|-----|--|----|-----|--|
| QY | Db | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | QY | 600 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | Db | 600 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | QY | 2247 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | Db | 2247 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | QY | 601 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | Db | 601 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | QY | 2307 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | Db | 2307 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | QY | 661 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | Db | 661 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | QY | 2367 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | Db | 2367 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | QY | 721 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | Db | 721 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | QY | 2427 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | Db | 2427 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | QY | 781 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | Db | 781 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | QY | 2487 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | Db | 2487 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | QY | 841 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | Db | 841 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | QY | 2547 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | Db | 2547 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | QY | 901 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | Db | 901 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | QY | 2607 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | Db | 2607 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | QY | 961 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC | Db | 961 | GGTGTCTATGACACAGGCGCATCCACACCATCGAGTTCTGCGCTGGGCTGGCTTCCAAACAC |

RESULT 8

```
US-11-266-748A-287153
; Sequence 287153, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 287153
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-287153
```

Query Match 36.9%; Score 973.2; DB 7; Length 1000;
Best Local Similarity 99.4%; Pred. No. 8.5e-166;
Matches 988; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

```
Query Match 36.9%; Score 973.2; DB 7; Length 1000;  
Best Local Similarity 99.4%; Pred. No. 8.5e-166;  
Matches 988; Conservative 0; Mismatches 3; Indels 3; Gaps 1;
```

| QY | Db | 1650 | 1710 | 61 | 1770 | 121 | 1830 | 181 | 1887 | 241 | 1947 | 301 | 2007 | 361 | 2067 | 421 | 2127 | 481 | 2187 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----|----|---|------|------|---|-----|------|---|------|------|---|-----|------|--|------|------|---|-----|------|--|----|------|---|----|------|--|----|-----|---|----|-----|--|----|------|---|----|------|--|----|------|---|----|------|--|----|-----|---|----|-----|--|----|------|---|----|------|--|----|-----|---|----|-----|---|
| QY | Db | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 1709 | QY | 1710 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 1709 | Db | 1710 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 60 | QY | 1770 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 1769 | Db | 1770 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 120 | QY | 1830 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 1829 | Db | 1830 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 180 | QY | 1887 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 1886 | Db | 1887 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 240 | QY | 241 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 1946 | Db | 241 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 300 | QY | 1947 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 2006 | Db | 1947 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 360 | QY | 2007 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 2066 | Db | 2007 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 420 | QY | 361 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 2126 | Db | 361 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 480 | QY | 2127 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 2186 | Db | 2127 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 540 | QY | 481 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 2246 | Db | 481 | CTCTTCAAGATGAAGATGTCGTATCTCTGGCGTGTGACATGCGCTTTGGGCTGT 2246 |

QY 1650 CTCCTTCAAGATGAAGATGTCCTCATCTGGGCGTGTGTCATGCGCTTTGGGGTGGT 1709
Db 1 CTCCTTCAAGATGAAGATGTCCTCATCTGGGCGTGTGTCATGCGCTTTGGGGTGGT 60
QY 1710 CCTCGAGTCTTCAACACGTCGACTTTGGCCAGAGGACCGGCTGTCTGGAGACGCT 1769
Db 61 CCTCGAGTCTTCAACACGTCGACTTTGGCCAGAGGACCGGCTGTCTGGAGACGCT 120
QY 1770 GCGGAGCTCACTTCTCTGCTGGGACTTTCGGTTACTCGTGTTCCTAGTCACTACAA 1829
Db 121 GCGGAGCTCACTTCTCTGCTGGGACTTTCGGTTACTCGTGTTCCTAGTCACTACAA 180
QY 1830 GTGGCTGTGTCTGGGCTGCCAGGCGGCGCTCG---CCAGCATCTCATCACTTCAT 1886
Db 181 GTGGCTGTGTCTGGGCTGCCAGGCGGCGCTCGGGCCCCCAGCATCTCATCACTTCAT 240
QY 1887 CAACATGTTCTCTTCTCCACAGCCCCAGCAACAGGCTGTCTACCCCCCGGAGGAGGT 1946
Db 241 CAACATGTTCTCTTCTCCACAGCCCCAGCAACAGGCTGTCTACCCCCCGGAGGAGGT 300
QY 1947 GTTCCAGGCAACGCTGTGTCTTGGCTTGGCCATGTTGCCATCTCTGCTGTGGCAC 2006
Db 301 GTTCCAGGCAACGCTGTGTCTTGGCTTGGCCATGTTGCCATCTCTGCTGTGGCAC 360
QY 2007 ACCCTCGACTGTGCACGCCACCGCGCGGCTGTGGAGAGGCGGCTGTACCCACA 2066
Db 361 ACCCTCGACTGTGCACGCCACCGCGCGGCTGTGGAGAGGCGGCTGTACCCACA 420
QY 2067 GGAGGAAAAAAGGCGGGTGTGTGGACCTGTGCGATCTGTGAATGGCTGGAGCTC 2126
Db 421 GGAGGAAAAAAGGCGGGTGTGTGGACCTGTGCGATCTGTGAATGGCTGGAGCTC 480
QY 2127 CGATGAGAAAAAGGCGGGCTTGGATGATGAAGAGAGGCGGAGCTGTCCCTCCGA 2186
Db 481 CGATGAGAAAAAGGCGGGCTTGGATGATGAAGAGAGGCGGAGCTGTCCCTCCGA 540
QY 2187 GTGTCTCATGCACCGCCATCCACACCATCGAGTTCTGTGCTGGGCTGTCTCCACAC 2246
Db 541 GTGTCTCATGCACCGCCATCCACACCATCGAGTTCTGTGCTGGGCTGTCTCCACAC 600
QY 2247 CGCTCTCTACTGTGGCTGTGGGCTGTAGCTGTGGCCACCGCCAGCTGTCCGAGGTCT 2306
Db 601 CGCTCTCTACTGTGGCTGTGGGCTGTAGCTGTGGCCACCGCCAGCTGTCCGAGGTCT 660
QY 2307 GTGGGCCATCGTGTATGCGCATAGGCTGTGGCTGTGGCGGAGGTGTGGCGGTGT 2366
Db 661 GTGGGCCATCGTGTATGCGCATAGGCTGTGGCTGTGGCGGAGGTGTGGCGGTGT 720
QY 2367 GTGTGCTGGTCCCATCTTTGCGGCTTTGCGGTGTGACCGTGTGCTATCTGTGTGTAT 2426
Db 721 GTGTGCTGGTCCCATCTTTGCGGCTTTGCGGTGTGACCGTGTGCTATCTGTGTGTAT 780
QY 2427 GGAGGAGTCTCTAGCTTCTGTGACGCCCTGTGGCTGTGCTGTGGAATTCAGAACAA 2486
Db 781 GGAGGAGTCTCTAGCTTCTGTGACGCCCTGTGGCTGTGCTGTGGAATTCAGAACAA 840
QY 2487 GTTCTACTCAGGCAACGCGCTTCAAGCTGTAGTCCCTTCACTTGTGCTGCCACAGTACTA 2546
Db 841 GTTCTACTCAGGCAACGCGCTTCAAGCTGTAGTCCCTTCACTTGTGCTGCCACAGTACTA 900
QY 2547 GGGCCCACTGACGCTCTGACAGCTCTTCTGTGCTGTGAGCTGTGAGGAGGAGGAGATAA 2606
Db 901 GGGCCCACTGACGCTCTGACAGCTCTTCTGTGCTGTGAGCTGTGAGGAGGAGGAGATAA 960
QY 2607 GACGGTCCGCTTGGCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
Db 961 GACGGTCCGCTTGGCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 994

RESULT 9

US-11-266-748A-338582/c

; Sequence 338582, Application US/11266748A

; Publication No. US2006013463A1

GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 5815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 338582
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-338582

Query Match 36.9%; Score 973.2; DB 7; Length 1000;

Best Local Similarity 99.4%; Pred. No. 8 Se-166;

Matches 988; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 1650 CTCCTTCAAGATGAAGATGTCCTCATCTGGGCGTGTGTCATGCGCTTTGGGGTGGT 1709
Db 1000 CTCCTTCAAGATGAAGATGTCCTCATCTGGGCGTGTGTCATGCGCTTTGGGGTGGT 941
QY 1710 CCTCGAGTCTTCAACACGTCGACTTTGGCCAGAGGACCGGCTGTCTGGAGACGCT 1769
Db 940 CCTCGAGTCTTCAACACGTCGACTTTGGCCAGAGGACCGGCTGTCTGGAGACGCT 881
QY 1770 GCGGAGCTCACCTTCTGTGGGACTTCTTCGGTTACTCTGTCTCTAGTCACTACAA 1829
Db 880 GCGGAGCTCACCTTCTGTGGGACTTCTTCGGTTACTCTGTCTCTAGTCACTACAA 821
QY 1830 GTGGCTGTGTCTGGGCTGCCAGGCGGCTCG---CCAGCATCTCATCACTTCAT 1886
Db 820 GTGGCTGTGTCTGGGCTGCCAGGCGGCTCGGGCCCCCAGCATCTCATCACTTCAT 761
QY 1887 CAACATGTTCTCTTCTCCACAGCCCCAGCAACAGGCTGTCTACCCCCCGGAGGAGGT 1946
Db 760 CAACATGTTCTCTTCTCCACAGCCCCAGCAACAGGCTGTCTACCCCCCGGAGGAGGT 701
QY 1947 GTTCCAGGCAACGCTGTGTCTTGGCCATGTTGCCATCTGTGCTGTGGCAC 2006
Db 700 GTTCCAGGCAACGCTGTGTCTTGGCCATGTTGCCATCTGTGCTGTGGCAC 641
QY 2007 ACCCTCGACTGTGCACCGCCACCGCGGCTGTGGAGAGGCGGCTGTGAGGAGGAGGAG 2066
Db 640 ACCCTCGACTGTGCACCGCCACCGCGGCTGTGGAGAGGCGGCTGTGAGGAGGAGGAG 581
QY 2067 GGAGGAAAAAAGGCGGGTGTGTGGACCTGTGCGATCTGTGAATGGCTGGAGCTC 2126
Db 580 GGAGGAAAAAAGGCGGGTGTGTGGACCTGTGCGATCTGTGAATGGCTGGAGCTC 521
QY 2127 CGATGAGAAAAAGGCGGGCTTGGATGATGAAGAGAGGAGGAGGAGGAGGAGGAGGAG 2186
Db 520 CGATGAGAAAAAGGCGGGCTTGGATGATGAAGAGAGGAGGAGGAGGAGGAGGAGGAG 461


```
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 468814
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
; US-11-266-748A-468814

Query Match      36.9%; Score 973.2; DB 7; Length 1000;
Best Local Similarity 99.4%; Pred. No. 8.5e-166;
Matches 988; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

Qy 1650 CTCCTTCAAGATGAAGATGTCGGTATCTCTGGGCGTCGTGCACATGCGCCCTTTGGGGTGT 1709
Db 1000 CTCCTTCAAGATGAAGATGTCGGTATCTCTGGGCGTCGTGCACATGCGCCCTTTGGGGTGT 941

Qy 1710 CCTCGGAGTCTTCAACACAGTGCATTTGGCCAGAGGACCGGCTGCTGTGAGAGCT 1769
Db 940 CCTCGGAGTCTTCAACACAGTGCATTTGGCCAGAGGACCGGCTGCTGTGAGAGCT 881

Qy 1770 GCCGGAGCTCACCTTCTCTGCTGGGACTCTTTCGGTTACTCTGTTTCTAGTCTATCAAA 1829
Db 880 GCCGGAGCTCACCTTCTCTGCTGGGACTCTTTCGGTTACTCTGTTTCTAGTCTATCAAA 821

Qy 1830 GTGGCTGTGTCTGGGCTGCCAGGGCGGCTCG---CCAGCATCTCTCATCCACTTTCAT 1886
Db 820 GTGGCTGTGTCTGGGCTGCCAGGGCGGCTCGGGCCCCCAGCATCTCTCATCCACTTTCAT 761

Qy 1887 CAACATGTTCTCTTCTCCACAGCCCCCAGCAGCTCTTACCCCCGGGAGAGGT 1946
Db 760 CAACATGTTCTCTTCTCCACAGCCCCCAGCAGCTCTTACCCCCGGGAGAGGT 701

Qy 1947 GTTCCAGGCCACGCTGTGTGTCTGGCTTTGGCCATGCTGCCATCTCTGCTGTGGCAC 2006
Db 700 GTTCCAGGCCACGCTGTGTGTCTGGCTTTGGCCATGCTGCCATCTCTGCTGTGGCAC 641

Qy 2007 ACCCTCTGCATCTGTGCACCGCCACCGCCGCTGCGGAGGAGGCGCCGCTGACCGACA 2066
Db 640 ACCCTCTGCATCTGTGCACCGCCACCGCCGCTGCGGAGGAGGCGCCGCTGACCGACA 581

Qy 2067 GGAGGAAAAAAGCCCGGGTGTGTGACCTGCTGAGCATCTGTGAATGGCTTGGAGCTC 2126
Db 580 GGAGGAAAAAAGCCCGGGTGTGTGAGCTGCTGAGCATCTGTGAATGGCTTGGAGCTC 521

Qy 2127 CGATGAGGAAAAAGCCAGGGGCTTGGATGATGATGAGGAGGCGGAGCTGCTCCCTCCGA 2186
Db 520 CGATGAGGAAAAAGCCAGGGGCTTGGATGATGATGAGGAGGCGGAGCTGCTCCCTCCGA 461
```

RESULT 12

```
US-11-266-748A-76628
; Sequence 76628, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76628
; LENGTH: 738
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45)..(45)
; OTHER INFORMATION: n is a, c, g, or t
```


| Result No. | Query | | | DB | ID | Description |
|------------|--------|-------|--------|----|---------------------|---------------------|
| | Score | Match | Length | | | |
| 1 | 2640 | 100.0 | 2640 | 10 | US-10-783-519-1 | Sequence 1, Appli |
| 2 | 2640 | 100.0 | 2655 | 3 | US-09-962-436-278 | Sequence 278, Appli |
| 3 | 2840 | 100.0 | 2655 | 3 | US-09-880-107-3363 | Sequence 3363, Ap |
| 4 | 2640 | 100.0 | 2655 | 3 | US-09-968-070-2192 | Sequence 212, App |
| 5 | 2640 | 100.0 | 2655 | 7 | US-10-305-072-1094 | Sequence 1094, Ap |
| 6 | 2640 | 100.0 | 2655 | 8 | US-10-641-643-916 | Sequence 916, App |
| 7 | 2640 | 100.0 | 2655 | 10 | US-10-843-641A-2737 | Sequence 2737, Ap |
| 8 | 2640 | 100.0 | 2655 | 10 | US-10-843-641A-6682 | Sequence 6682, Ap |
| 9 | 2840 | 100.0 | 2655 | 16 | US-11-122-329-72 | Sequence 72, Appli |
| 10 | 2068 | 78.3 | 2488 | 7 | US-10-145-012-1 | Sequence 1, Appli |
| 11 | 2068 | 78.3 | 2488 | 13 | US-11-126-866-1 | Sequence 1, Appli |
| 12 | 2068 | 78.3 | 2488 | 15 | US-11-126-841A-1 | Sequence 1, Appli |
| 13 | 2066.4 | 78.3 | 2488 | 7 | US-10-145-012-12 | Sequence 12, Appli |
| 14 | 2066.4 | 78.3 | 2488 | 13 | US-11-126-866-12 | Sequence 12, Appli |
| 15 | 2066.4 | 78.3 | 2488 | 15 | US-11-126-841A-12 | Sequence 12, Appli |
| 16 | 1989.4 | 75.4 | 3104 | 7 | US-10-264-049-481 | Sequence 481, App |
| 17 | 742.8 | 28.1 | 2937 | 10 | US-10-764-420-2037 | Sequence 2037, Ap |

181 GCTCGGTGAGCGCTTCAGAGACGCTTTGTGTTGATGTTTGGCGCTGTGAGAGCTG 240
241 GAGAGACCTTCACTTCTCTGAGAGAGAGGTGCGCGGGCTGGGTGCTTGCCTCCCG 300
241 GAGAGACCTTCACTTCTCTGAGAGAGAGGTGCGCGGGCTGGGTGCTTGCCTCCCG 300
301 CAAAGGGAGGCTGCCGACCCACCCCGGACCTGCTGCGCATCCAGAGGAGAGCG 360
301 CAAAGGGAGGCTGCCGACCCACCCCGGACCTGCTGCGCATCCAGAGGAGAGCG 360
361 GAGCGCTTGGCCAGAGCTGCGGGATGTGCGGGCAACAGAGGCGCTTGCGGGCCAG 420
361 GAGCGCTTGGCCAGAGCTGCGGGATGTGCGGGCAACAGAGGCGCTTGCGGGCCAG 420
421 CTGACACAGCTGAGCTCAGCGCGCGCTGCTACGCGCGGCGCATGAACCTCAGCTGGCA 480
421 CTGACACAGCTGAGCTCAGCGCGCGCTGCTACGCGCGGCGCATGAACCTCAGCTGGCA 480
481 GCCGCCACACAGATGGGGCTCAGAGAGAGCGCCCTGCTCCAGGCGCCCGGGGGCG 540
481 GCCGCCACACAGATGGGGCTCAGAGAGAGCGCCCTGCTCCAGGCGCCCGGGGGCG 540
541 CACAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGGTGGAGCGCCCAAGGCGCCCTGCC 600
541 CACAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGGTGGAGCGCCCAAGGCGCCCTGCC 600
601 CTAGAGCGCTGCTCTGGAGGGCTGCGCGGCTTCTCATTTGCCAGCTTCAGGAGCTG 660
601 CTAGAGCGCTGCTCTGGAGGGCTGCGCGGCTTCTCATTTGCCAGCTTCAGGAGCTG 660
661 GAGCAGCGCTGAGCAGACCCCGTGAGCGGCGAGCGCAGCTGAGTGAATCTTCCTCATC 720
661 GAGCAGCGCTGAGCAGACCCCGTGAGCGGCGAGCGCAGCTGAGTGAATCTTCCTCATC 720
721 TCTACTGGGGTGAAGATCGGACAGAGATCCGCAAGATCAGCGACTGCTTCCACTGC 780
721 TCTACTGGGGTGAAGATCGGACAGAGATCCGCAAGATCAGCGACTGCTTCCACTGC 780
781 CAGCTTCTCCGTTCTGACAGAGAGAGGCGCGCTCGGGCCCTGACAGCTGCAAA 840
781 CAGCTTCTCCGTTCTGACAGAGAGAGGCGCGCTCGGGCCCTGACAGCTGCAAA 840
841 CAGCAGCAGGAGCTGAGAGGTCTCGGGGAGACAGAGCGGTCTCTGAGCAGAGTG 900
841 CAGCAGCAGGAGCTGAGAGGTCTCGGGGAGACAGAGCGGTCTCTGAGCAGAGTG 900
901 CTAGGCGGGTGTGAGCTGTGCGCGCAGGCGAGGTGCAAGTCCAAAGATGAAGGCC 960
901 CTAGGCGGGTGTGAGCTGTGCGCGCAGGCGAGGTGCAAGTCCAAAGATGAAGGCC 960
961 GTGTACTGGCTGAACAGGTGAGCGTGAGCACCACCAAGTGCCTCATTTGCCGAG 1020
961 GTGTACTGGCTGAACAGGTGAGCGTGAGCACCACCAAGTGCCTCATTTGCCGAG 1020
1021 GCCTGTGTCTGTGCGAGACTGCGCGCTGACAGGCGCTCGGGGACAGCTCGATG 1080
1021 GCCTGTGTCTGTGCGAGACTGCGCGCTGACAGGAGGCGCTCGGGGACAGCTCGATG 1080
1081 GAGGAGGAGTGAAGTCCGTGCTACCGCATCCCTGCGGGACATGCCGCCCAACATC 1140
1081 GAGGAGGAGTGAAGTCCGTGCTACCGCATCCCTGCGGGACATGCCGCCCAACATC 1140
1141 ATCCGACCAACCGCTTACGGCGAGCTTCCAGGCGATCGTGGATCGCTACGGGTGGGC 1200
1141 ATCCGACCAACCGCTTACGGCGAGCTTCCAGGCGATCGTGGATCGCTACGGGTGGGC 1200
1201 CGCTACAGAGGTCAACCCCGCTTACACCATCATCACCTTCCCTTCTGTTGCT 1260
1201 CGCTACAGAGGTCAACCCCGCTTACACCATCATCACCTTCCCTTCTGTTGCT 1260
1261 GTGATGTTGGGGATGGGCCACCGGCTGCTCATGTTCTCTTGCCTTGGCCATGGTC 1320

1261 GTGATGTTGGGGATGGGCCACCGGCTGCTCATGTTCTCTTGCCTTGGCCATGGTC 1320
1321 CTTTGGGAGAACCGACCGCTGTGAAAGCCGCGAGAACGAGATCTGGCAGACTTCTTTC 1380
1321 CTTTGGGAGAACCGACCGCTGTGAAAGCCGCGAGAACGAGATCTGGCAGACTTCTTTC 1380
1381 AGGGGCGGTACTGCTCTGCTGTTTGGGCTGTTCTCCATCTACACCGGCTTCATCTAC 1440
1381 AGGGGCGGTACTGCTCTGCTGTTTGGGCTGTTCTCCATCTACACCGGCTTCATCTAC 1440
1441 AACAGTGTCTCAGTGGGCCACAGCATTTCCCTCGGGCTGGAGTGTGGCGGCAATG 1500
1441 AACAGTGTCTCAGTGGGCCACAGCATTTCCCTCGGGCTGGAGTGTGGCGGCAATG 1500
1501 GCCAACAGTCTGGCTGGAGTGTGATGATCTTCCGCGGCTGGAGTGTGGCGGCAATG 1560
1501 GCCAACAGTCTGGCTGGAGTGTGATGATCTTCCGCGGCGGCAACAGATGCTTACCTGGAT 1560
1561 CCCAACAGTCTGGCTGGAGTGTGATGATCTTCCGCGGCGGCAACAGATGCTTACCTGGAT 1620
1561 CCCAACAGTCTGGCTGGAGTGTGATGATCTTCCGCGGCGGCAACAGATGCTTACCTGGAT 1620
1621 CTGGCTGCGCAACCACTTGGCTTCTCAACTCTTCAAGATGAAGATGTCGCTCATCTTC 1680
1621 CTGGCTGCGCAACCACTTGGCTTCTCAACTCTTCAAGATGAAGATGTCGCTCATCTTC 1680
1681 GCGCTGCTGCAATGGCTTGGGCTGCTTCCGAGTCTTCAACAGTGCACCTTTGGC 1740
1681 GCGCTGCTGCAATGGCTTGGGCTGCTTCCGAGTCTTCAACAGTGCACCTTTGGC 1740
1741 CAGAGCACCGGCTGCTGCTGAGACGCTGCGGAGCTCACCCTTCTGCTGGGACTTCTTC 1800
1741 CAGAGCACCGGCTGCTGCTGAGACGCTGCGGAGCTCACCCTTCTGCTGGGACTTCTTC 1800
1801 GGTACTCTGTTCTAGTCTATCAAGTGGCTGTGTGTCTGGCTGCTGAGGCGCGC 1860
1801 GGTACTCTGTTCTAGTCTATCAAGTGGCTGTGTGTCTGGCTGCTGAGGCGCGC 1860
1861 TCGCCGAGGACTCTATCAGCTTCAATGATGTTCTTCTTCCACAGGCGGCGGCAAC 1920
1861 TCGCCGAGGACTCTATCAGCTTCAATGATGTTCTTCTTCCACAGGCGGCGGCAAC 1920
1921 AGGCTGCTTACCCCGGAGAGGTGGTCCAGGCGCACGCTGCTGGTCTTGGCTTGGGC 1980
1921 AGGCTGCTTACCCCGGAGAGGTGGTCCAGGCGCACGCTGCTGGTCTTGGCTTGGGC 1980
1981 ATGGTGGCCATCTGCTGCTGGCACAACCTTGCACCTGCTGCAACGCGCGCGCGC 2040
1981 ATGGTGGCCATCTGCTGCTGGCACAACCTTGCACCTGCTGCAACGCGCGCGCGC 2040
2041 CTTGGGAGGAGGCGCGCTGACCGAGGAGGAAACAGGCGGGTGTGTGGACCTGCTTC 2100
2041 CTTGGGAGGAGGCGCGCTGACCGAGGAGGAAACAGGCGGGTGTGTGGACCTGCTTC 2100
2101 GACGATCTGTGAATGGCTGAGCTCCGATGAGGAAAGAGGCGGCTTGGATGATGA 2160
2101 GACGATCTGTGAATGGCTGAGCTCCGATGAGGAAAGAGGCGGCTTGGATGATGA 2160
2161 GAGGAGGCGAGCTGCTGCTTCCAGGCTGCTGATGAGGCGGCTTGGATGATGA 2220
2161 GAGGAGGCGAGCTGCTGCTTCCAGGCTGCTGATGAGGCGGCTTGGATGATGA 2220
2221 TTTGCTGGCTGGCTGCTTCCAGCAGGCGCTTCTGCTGCGCTTGGGCGCTGAGGCTG 2280
2221 TTTGCTGGCTGGCTGCTTCCAGCAGGCGCTTCTGCTGCGCTTGGGCGCTGAGGCTG 2280
2281 GCGCCAGCGGCTGCTGCTGAGGCTTCTGCTGGGCGATGATGCGATGAGGCGCTG 2340
2281 GCGCCAGCGGCTGCTGCTGAGGCTTCTGCTGGGCGATGATGCGATGAGGCGCTG 2340
2341 GCGCGGAGGCTGGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
2341 GCGCGGAGGCTGGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400

QY 2401 ATGACCGTGGCTATCTCTGCTGGTGTATGAGAGGAACTCTCAGCCTTCTGCAAGCCCTGCGG 2460
DB 2401 ATGACCGTGGCTATCTCTGCTGGTGTATGAGAGGAACTCTCAGCCTTCTGCAAGCCCTGCGG 2460
QY 2461 CTGCACTGGGTGGAAATTCACAGAACAGTCTTCTACTCAGGCAAGGCTTACAGCTGAGTCCC 2520
DB 2461 CTGCACTGGGTGGAAATTCACAGAACAGTCTTCTACTCAGGCAAGGCTTACAGCTGAGTCCC 2520
QY 2521 TTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
DB 2521 TTCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
QY 2581 GACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
DB 2581 GACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640

RESULT 2
US-09-962-436-278
; Sequence 278, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; PRIORITY FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIORITY FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIORITY FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 278
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-278

Query Match 100.0%; Score 2640; DB 3; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCTGCGCGAGCGGACAGCCAGCAGCGAGGCGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAG 60
DB 1 CGCGCTGCGCGAGCGGACAGCCAGCAGCGAGGCGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAG 60
QY 61 GGCCTCATGTTCCGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
DB 61 GGCCTCATGTTCCGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY 121 GCCTACACCTGCTGAGTGGCTGGGCGAGCTGGGCTGGTGGAGTTTCAGAGACCTCAAC 180
DB 121 GCCTACACCTGCTGAGTGGCTGGGCGAGCTGGGCTGGTGGAGTTTCAGAGACCTCAAC 180
QY 181 GCCTCGGTGAGCGCTTCCAGAGACGCTTTGTTGTTGATTTTGGCGCTGTGAGGAGCTG 240
DB 181 GCCTCGGTGAGCGCTTCCAGAGACGCTTTGTTGTTGATTTTGGCGCTGTGAGGAGCTG 240
QY 241 GAGAAACCTTACCTTCTGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
DB 241 GAGAAACCTTACCTTCTGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
QY 301 CCAAGAG 360
DB 301 CCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
QY 361 GAGCGCTGCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
DB 361 GAGCGCTGCGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420

QY 421 CTGCAACAGCTGAGAGCTTCAACGCGCGCTGCTAGCCAGGCGCCATGAACCTCAGCTGCA 480
DB 421 CTGCAACAGCTGAGAGCTTCAACGCGCGCTGCTAGCCAGGCGCCATGAACCTCAGCTGCA 480
QY 481 GCGCCGACACAGATGGGGCTCAGAGAGAGCGCCCTGCTCAGGCGCCCGGGGGGGCG 540
DB 481 GCGCCGACACAGATGGGGCTCAGAGAGAGCGCCCTGCTCAGGCGCCCGGGGGGGCG 540
QY 541 CACAGAGACCTGAGGGTCAACTTTGTGGCAGGTGCGCTGGAGCCCAAGAGCCCTGCGC 600
DB 541 CACAGAGACCTGAGGGTCAACTTTGTGGCAGGTGCGCTGGAGCCCAAGAGCCCTGCGC 600
QY 601 CTAGAGCGCTGCTGAGAGGCGCTGCGCGGGCTTCTCATTTGCGAGCTTTCAGGAGCTG 660
DB 601 CTAGAGCGCTGCTGAGAGGCGCTGCGCGGGCTTCTCATTTGCGAGCTTTCAGGAGCTG 660
QY 661 GAGCAGCGCTGAGAGCAACCCGTGACGGGCGAGCAGCAGCAGCTGGATGACCTTCTCATC 720
DB 661 GAGCAGCGCTGAGAGCAACCCGTGACGGGCGAGCAGCAGCAGCTGGATGACCTTCTCATC 720
QY 721 TCCTACTGGGGTGAAGAGATCGGCAAGATCGCAAGATCAGGACTGCTTCCACTGC 780
DB 721 TCCTACTGGGGTGAAGAGATCGGCAAGATCGGCAAGATCAGGACTGCTTCCACTGC 780
QY 781 CACGTCTTCCGCTTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
DB 781 CACGTCTTCCGCTTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
QY 841 CAGCAGAGCAGGAGCTGAGGAGGAGTCTCGGGGAGACAGAGGAGTTCCTGAGGCGAGTG 900
DB 841 CAGCAGAGCAGGAGCTGAGGAGGAGTCTCGGGGAGACAGAGGAGTTCCTGAGGCGAGTG 900
QY 901 CTAGGCGGGTGTGAGCTGCTGCGCCAGGCGAGGAGTGCAGGTCCACAGATGAAGGCC 960
DB 901 CTAGGCGGGTGTGAGCTGCTGCGCCAGGCGAGGAGTGCAGGTCCACAGATGAAGGCC 960
QY 961 GTGTACTGGCCCTGAGACAGCTGAGGAGCAGCAGCAGCTGATGCTCATTTGCCGAG 1020
DB 961 GTGTACTGGCCCTGAGACAGCTGAGGAGCAGCAGCAGCTGATGCTCATTTGCCGAG 1020
QY 1021 GCCTGTGCTCTGTCGAGAGCTGCGCGCCCTGCGAGGAGGCGCTGCGGAGACGCTCGATG 1080
DB 1021 GCCTGTGCTCTGTCGAGAGCTGCGCGCCCTGCGAGGAGGCGCTGCGGAGACGCTCGATG 1080
QY 1081 GAGGAGGAGTGTAGTGTGCTGCGCTCAGCGCATCCCTGCGGAGACATGCCCCCAGCATC 1140
DB 1081 GAGGAGGAGTGTAGTGTGCTGCGCTCAGCGCATCCCTGCGGAGACATGCCCCCAGCATC 1140
QY 1141 ATCCGACACCAACCGCTTACGGCGAGTTCAGGGGATCGTGGATCGCTACGGCGTGGG 1200
DB 1141 ATCCGACACCAACCGCTTACGGCGAGTTCAGGGGATCGTGGATCGCTACGGCGTGGG 1200
QY 1201 CGCTACAGAGGTCAACCCCGCTCCCTACACCATCATACCTTCCCTTCCCTTCTGTTGCT 1260
DB 1201 CGCTACAGAGGTCAACCCCGCTCCCTACACCATCATACCTTCCCTTCCCTTCTGTTGCT 1260
QY 1261 GTGATGTTGCGGAGTGTGGCCACAGGCTGCTCATGTTTCTTTCGCGCTGGCCATGGTC 1320
DB 1261 GTGATGTTGCGGAGTGTGGCCACAGGCTGCTCATGTTTCTTTCGCGCTGGCCATGGTC 1320
QY 1321 CTTGCGGAGAACCGGCTGTGAAAGCGCGCAGAACAGATCTGGGAGACTTCTTTC 1380
DB 1321 CTTGCGGAGAACCGGCTGTGAAAGCGCGCAGAACAGATCTGGGAGACTTCTTCTTTC 1380
QY 1381 AGGGGCGCTTACCTGCTGCTGTTATGGGCTGTTCTTCAATCATACCTTCCCTTCTGTTGCT 1440
DB 1381 AGGGGCGCTTACCTGCTGCTGTTATGGGCTGTTCTTCAATCATACCTTCCCTTCTGTTGCT 1440
QY 1441 AACAGTGTCTTCACTGCGCGCACAGCATCTTCCCTCGGGCTGGAGTGTGGCGCCATG 1500
DB 1441 AACAGTGTCTTCACTGCGCGCACAGCATCTTCCCTCGGGCTGGAGTGTGGCGCCATG 1500
QY 1501 GCCAACAGTCTGGCTGGAGTGTGATGATTCCTGGCGCCAGCACAGATGCTTACCTGGAT 1560

| | | | |
|----|------|---|------|
| Qy | 601 | CTAGAGCGCCTGCTCTGAGAGCGCTGCGCGGCTTCCTCATTTGCCAGCTTCAGGAGCTG | 660 |
| Ds | 601 | | 660 |
| Qy | 661 | GAGCAGCGCTGAGAGCAACCCGCTGACCGGCGGAGCCAGCCACGCTGGATGACCTTCCTCATC | 720 |
| Ds | 661 | | 720 |
| Qy | 721 | TCCTACTGGGGTGAAGAGATCGGACAGAAAGATCCGCAAGATCAAGGATCGCTTCCTCATC | 780 |
| Ds | 721 | | 780 |
| Qy | 781 | CACGCTCTCCGCTTCTGACAGAGAGAGAGCGCGCTCGGGGCGCTCGCAGAGTCGAA | 840 |
| Ds | 781 | | 840 |
| Qy | 841 | CAGCAGAGCCAGGAGCTGCTCGGGGAGACAGAGCGGTTCTCGAGCCAGGTG | 900 |
| Ds | 841 | | 900 |
| Qy | 901 | CTAGGCGGGTGTGAGCTGCTGCGCCAGGCGAGGTGCAAGATGAAGGCC | 960 |
| Ds | 901 | | 960 |
| Qy | 961 | GTGTACCTGGCCCTGAACCACTGAGCTGAGACACGACCAAGTGCCTCATTTGCCGAG | 1020 |
| Ds | 961 | | 1020 |
| Qy | 1021 | GCTGTGTGCTGTGCGGAGACTGCGCGCCCTGCGAGGAGCGCTCGGGGAGAGCTCGATG | 1080 |
| Ds | 1021 | | 1080 |
| Qy | 1081 | GAGAGGAGTGAAGTGGCTGCTACCGCATATCATCTTCCCTGCGGGGACATGCCCCCACATC | 1140 |
| Ds | 1141 | | 1140 |
| Qy | 1141 | ATCGCACCAACCGCTTACGGCCAGCTTCAGGGGATCGTGATCGTACGGGTGGGC | 1200 |
| Ds | 1141 | | 1200 |
| Qy | 1201 | CGCTACAGGAGGTCAACCCCGCTCCCTACACCATCATCTTCCCTTCTCTTTGCT | 1260 |
| Ds | 1201 | | 1260 |
| Qy | 1261 | GTGATGTTGGGGATGTGGGCCACGGGCTGCTCATGTTCTTCTTGGCCCTGGCCATGGTC | 1320 |
| Ds | 1261 | | 1320 |
| Qy | 1321 | CTTGGGAGAACCGACCGGCTGTGAAGCGCGGAGAACGAGATCTGGCAGACTTTCTTC | 1380 |
| Ds | 1321 | | 1380 |
| Qy | 1381 | AGGGCGGCTACCTGCTCTGCTTATGGGCTGTCTTCCATCTACACCGGCTTCATCTAC | 1440 |
| Ds | 1381 | | 1440 |
| Qy | 1441 | AACGAGTGTTCAGTTCGCGCACAGCATCTTCCCTCGGGCTGGAGTGGCGGCCATG | 1500 |
| Ds | 1441 | | 1500 |
| Qy | 1501 | GCCAACGAGTCTGGCTGAGTATGATTCCTGGGCCAGCAGCATGCTTACCTGGAT | 1560 |
| Ds | 1501 | | 1560 |
| Qy | 1561 | CCCAAGCTACCGGTGTCTTCTGGGACCTTACCCTTTGGCATCGATCTATTGGAGC | 1620 |
| Ds | 1561 | | 1620 |
| Qy | 1621 | CTGGCTGCCAACCACTTGAAGTCTCTCAACTCTTCAAGATGAAGATGTCGCTCATCTG | 1680 |
| Ds | 1621 | | 1680 |

| | | | |
|----|------|--|------|
| Qy | 1681 | GGCGTCTGTGCATATGGCCCTTTGGGGTGTCTCTCGAGTCTTCAACCAAGTGCATTTGGC | 1740 |
| Ds | 1681 | | 1740 |
| Qy | 1741 | CAGAGGACACCGGCTGTCTGTGAGACGCTGCGGAGCTCACCTTCTCTGGGACTCTTC | 1800 |
| Ds | 1741 | | 1800 |
| Qy | 1801 | GGTTACCTCTGTGTTCTTAGTCACTTACAAGTGGTGTGTCTGGGCTGCCAGGCCGCC | 1860 |
| Ds | 1801 | | 1860 |
| Qy | 1861 | TCGCCGAGACTCTCATCTTCAATCAAGTGTCTTCTTCTCCACAGCCCCAGCAAC | 1920 |
| Ds | 1861 | | 1920 |
| Qy | 1921 | AGGCTGTCTTACCCCGGAGAGGTGTCCAGGCAACGCTGTGTCTGGCTTGGCC | 1980 |
| Ds | 1921 | | 1980 |
| Qy | 1981 | ATGTGCGCCATCTGTCTTGGCACACCTCTGCACTGTCTGCAACCGCCACCGCCGCCG | 2040 |
| Ds | 1981 | | 2040 |
| Qy | 2041 | CTGCGGAGAGCCCGCTGACCGGACAGGAGGAAAACAAAGGCGGGTCTGTGACCTGCT | 2100 |
| Ds | 2041 | | 2100 |
| Qy | 2101 | GACGCATCTGTGAATGTGAGCTCCGATGAGGAAAAGGAGGGGCTCGATGATGAA | 2160 |
| Ds | 2101 | | 2160 |
| Qy | 2161 | GAGAGGCCGAGTCTGCTCCCTCCGAGTGTCTATGCAACAGGCGCATCACACCATCGAG | 2220 |
| Ds | 2161 | | 2220 |
| Qy | 2221 | TTCTGCTGGGCTGGCTCTCCAAACACCGCTCTACCTGCGCTGTGGGCGCTTGAGCCTG | 2280 |
| Ds | 2221 | | 2280 |
| Qy | 2281 | GCCCAACCGGAGTCTCGAGGTTCTGTGGGCAATGTGTGCGCATAGGCTTGGGCTG | 2340 |
| Ds | 2281 | | 2340 |
| Qy | 2341 | GGCGGAGGTGGCGGT | 2400 |
| Ds | 2341 | | 2400 |
| Qy | 2401 | ATGACCGTGGCTATCTCTGT | 2460 |
| Ds | 2401 | | 2460 |
| Qy | 2461 | CTGCATGGGTGGAAATTCAGAACAAAGTCTTACTCAGGCGGGCTACAGCTGAGTCCC | 2520 |
| Ds | 2461 | | 2520 |
| Qy | 2521 | TTCACTTCTGTCACAGATGACTAGGGGCCCACTGCGAGTCTTCTGCCAGCTCTCTTCT | 2580 |
| Ds | 2521 | | 2580 |
| Qy | 2581 | GACCTGTGAGGAGGAGGAGGAAAGACGGTCCGCCCTTGGCAAAAAAAGAAAAA | 2640 |
| Ds | 2581 | | 2640 |

RESULT 4

US-09-968-007A-212
; Sequence 212, Application US/09968007A
; Publication No. US20040115625A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; TITLE OF INVENTION: Gene Sets

[illegible]

| | | | |
|----|------|---|------|
| Qy | 1801 | GTTTACCTCGTGTTCTTAGTCATCTACAAGTGGCTGTGTGTCCTGGGCTGCAGGCGCCG | 1860 |
| Dz | 1801 | GTTTACCTCGTGTTCTTAGTCATCTACAAGTGGCTGTGTGTCCTGGGCTGCAGGCGCCG | 1860 |
| Qy | 1861 | TGCGCCAGCATCCTCATCCACTTTCATCAACATGTTTCCTTCTCCCACAGCCCCAGCAAC | 1920 |
| Dz | 1861 | TGCGCCAGCATCCTCATCCACTTTCATCAACATGTTTCCTTCTCCCACAGCCCCAGCAAC | 1920 |
| Qy | 1921 | AGGCTGCTCTACCCCGCGCAGAGTGTGCCAGGCCACGCTGCTGGTCTGGCTTGGCC | 1980 |
| Dz | 1921 | AGGCTGCTCTACCCCGCGCAGAGTGTGCCAGGCCACGCTGCTGGTCTGGCTTGGCC | 1980 |
| Qy | 1981 | ATGTGGCCAPCTCTCTCTGCTGGCACACCCTTGCACTGCTGCAACCGCCACCGCGCCG | 2040 |
| Dz | 1981 | ATGTGGCCAPCTCTCTCTGCTGGCACACCCTTGCACTGCTGCAACCGCCACCGCGCCG | 2040 |
| Qy | 2041 | CTGCGGAGGCGCCGCTGACCGACAGGAGAAACAAGCCGGTTCCTGGACCTTGCT | 2100 |
| Dz | 2041 | CTGCGGAGGCGCCGCTGACCGACAGGAGAAACAAGCCGGTTCCTGGACCTTGCT | 2100 |
| Qy | 2101 | GACGCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAAAGGCAGGGGCGCTTGGATGATAA | 2160 |
| Dz | 2101 | GACGCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAAAGGCAGGGGCGCTTGGATGATAA | 2160 |
| Qy | 2161 | GAGAGGCGAGCTGCTGCCCTCCGAGGTGTCTATGCAACGAGGCATCCACCATCGAG | 2220 |
| Dz | 2161 | GAGAGGCGAGCTGCTGCCCTCCGAGGTGTCTATGCAACGAGGCATCCACCATCGAG | 2220 |
| Qy | 2221 | TTCTGCTGGGCTGCGCTCTCCAACACCGCTCTTACCTGGCGCTGCGGGCGCTGAGCGCTG | 2280 |
| Dz | 2221 | TTCTGCTGGGCTGCGCTCTCCAACACCGCTCTTACCTGGCGCTGCGGGCGCTGAGCGCTG | 2280 |
| Qy | 2281 | GCCCAACGCCAGCTGTCCGAGGTTCTGTGGGCCATATGTGATGCGCATAGGCCCTGGGCTG | 2340 |
| Dz | 2281 | GCCCAACGCCAGCTGTCCGAGGTTCTGTGGGCCATATGTGATGCGCATAGGCCCTGGGCTG | 2340 |
| Qy | 2341 | GGCCGGAGGTGGGCGTGGGCGCTGGTGTCCCATCTTTGGCGCTTTCGGCTG | 2400 |
| Dz | 2341 | GGCCGGAGGTGGGCGTGGGCGCTGGTGTCCCATCTTTGGCGCTTTCGGCTG | 2400 |
| Qy | 2401 | ATGACCGTGGCTATCTCTGCTGGTGAATGGAGGAGCTCTCAGCCCTTCCTGCAAGCCCTGCGG | 2460 |
| Dz | 2401 | ATGACCGTGGCTATCTCTGCTGGTGAATGGAGGAGCTCTCAGCCCTTCCTGCAAGCCCTGCGG | 2460 |
| Qy | 2461 | CTGCACTGGGTGGAAATTCCAGAACAAATTTCTATCAGGCAAGGCTTAAGAAGTGAATGCC | 2520 |
| Dz | 2461 | CTGCACTGGGTGGAAATTCCAGAACAAATTTCTACTCAGGCAAGGCTTAAGAAGTGAATGCC | 2520 |
| Qy | 2521 | TTCACTTTCGTCGCCACAGATGACTAGGGGCCACTGCAAGTCTCTGCGCAGACCTCTTCTCT | 2580 |
| Dz | 2521 | TTCACTTTCGTCGCCACAGATGACTAGGGGCCACTGCAAGTCTCTGCGCAGACCTCTTCTCT | 2580 |
| Qy | 2581 | GACCTCTGAGGAGGAGGAATAAGACGGTCCGCCCTGGCAAAAAAAAAAAAAAAAAAAAA | 2640 |
| Dz | 2581 | GACCTCTGAGGAGGAGGAATAAGACGGTCCGCCCTGGCAAAAAAAAAAAAAAAAAAAAA | 2640 |

RESULT 5

```

US-10-305-720-1094
; Sequence 1094, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1094

```


1861 TCGCCAGCATCTCATCCACTTCAACATGTTCTCTTCCACAGCCCGCAAC 1920
1861 TCGCCAGCATCTCATCCACTTCAACATGTTCTCTTCCACAGCCCGCAAC 1920
1921 AGGCTGCTTACCCCGGAGAGGTTGTCAGGCCACGCTGTGTCTCTGCTTGGC 1980
1921 AGGCTGCTTACCCCGGAGAGGTTGTCAGGCCACGCTGTGTCTCTGCTTGGC 1980
1981 ATGTTGCCATCTGCTGTCTTGGCACACCCCTGCACTGCTGCACCGCCACCGCCGC 2040
1981 ATGTTGCCATCTGCTGTCTTGGCACACCCCTGCACTGCTGCACCGCCACCGCCGC 2040
2041 CTGGGAGAGCCCGCTGACCGACAGGAGGAAACAAGGCCGGGTTGCTGACCTGCT 2100
2041 CTGGGAGAGCCCGCTGACCGACAGGAGGAAACAAGGCCGGGTTGCTGACCTGCT 2100
2101 GACGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGCAGGGGCTGATGATGAA 2160
2101 GACGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGCAGGGGCTGATGATGAA 2160
2161 GAGGAGCCGAGCTGCTCCCTCGAGGTTGTCATGCAACAGGCCATCCACCATCGAG 2220
2161 GAGGAGCCGAGCTGCTCCCTCGAGGTTGTCATGCAACAGGCCATCCACCATCGAG 2220
2221 TTCTGCTGGGCTGGGCTTCCACACCGGCTCTTACCTGGGCTGTGGGCTGAGCCTG 2280
2221 TTCTGCTGGGCTGGGCTTCCACACCGGCTCTTACCTGGGCTGTGGGCTGAGCCTG 2280
2281 GCCACGCCGAGCTGCTCCAGGTTCTGTGGGCTGATGCGATAGGCTGAGGCTG 2340
2281 GCCACGCCGAGCTGCTCCAGGTTCTGTGGGCTGATGCGATAGGCTGAGGCTG 2340
2341 GGCAGGAGGTTGGGCTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 2400
2341 GGCAGGAGGTTGGGCTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 2400
2401 ATGACGCTGGCTATCTGCTGGTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 2460
2401 ATGACGCTGGCTATCTGCTGGTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 2460
2461 CTGCACTGGTGGATTCAGAACAAAGTTCTACTCAGGCACGGGCTACAGCTGAGTCC 2520
2461 CTGCACTGGTGGATTCAGAACAAAGTTCTACTCAGGCACGGGCTACAGCTGAGTCC 2520
2521 TTCACCTTGGTGGCAGATGACTAGGCCCCACTGAGGCTCTGCGCAGACTCTCTTCT 2580
2521 TTCACCTTGGTGGCAGATGACTAGGCCCCACTGAGGCTCTGCGCAGACTCTCTTCT 2580
2581 GACCTCTGAGGAGGAGGATTAAGACGGTCCGCTGGCAAAAAA 2640
2581 GACCTCTGAGGAGGAGGATTAAGACGGTCCGCTGGCAAAAAA 2640

RESULT 7

US-10-843-641A-2737
Sequence 2737, Application US/10843641A
Publication No. US2005006454A1
GENERAL INFORMATION:
APPLICANT: Avalon Pharmaceuticals, Inc.
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
FILE REFERENCE: 684290-189
CURRENT APPLICATION NUMBER: US/10/843,641A
CURRENT FILING DATE: 2004-05-12
PRIORITY APPLICATION NUMBER: US/09/873,367
PRIORITY FILING DATE: 2001-06-05
PRIORITY APPLICATION NUMBER: US/09/954,531
PRIORITY FILING DATE: 2001-09-18
PRIORITY APPLICATION NUMBER: US/09/954,456
PRIORITY FILING DATE: 2001-09-25
PRIORITY APPLICATION NUMBER: US/09/962,436
PRIORITY FILING DATE: 2001-09-25

PRIOR APPLICATION NUMBER: US/09/962,832
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/964,824
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US/09/967,768
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US/09/968,007
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US/09/969,347
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US/09/969,708
PRIOR FILING DATE: 2001-10-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 8447
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2737
LENGTH: 2655
TYPE: DNA
ORGANISM: Homo sapiens
US-10-843-641A-2737

Query Match 100.0%; Score 2640; DB 10; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCTGCGCGACGGGACGCGAGCGGAGCGCGGCGAGCACACCCGGGGACCATG 60
DB 1 CGGCTGCGCGACGGGACGCGAGCGGAGCGCGGCGAGCACACCCGGGGACCATG 60

QY 61 GGCTCCATGTTCCGAGCGAGGAGTGGCCCTGGTGGTCTTCTTTCGCCACAGCGCT 120
DB 61 GGCTCCATGTTCCGAGCGAGGAGTGGCCCTGGTGGTCTTCTTTCGCCACAGCGCT 120

QY 121 GCCTACACTGCTGAGTGGCTGGGCGAGCTGGGCTCTGTGGAGTTTCAGAGACCTCAAC 180
DB 121 GCCTACACTGCTGAGTGGCTGGGCGAGCTGGGCTCTGTGGAGTTTCAGAGACCTCAAC 180

QY 181 GCCTGGTGGGCGCTTCAGAGAGCGCTTGTGGTGTGATGTTGGCGCTGTGAGAGCTG 240
DB 181 GCCTGGTGGGCGCTTCAGAGAGCGCTTGTGGTGTGATGTTGGCGCTGTGAGAGCTG 240

QY 241 GAGAGAGCTTCACTCTCTCTGAGGAGGAGTGGGCGGCTGGGCTGTGCTCTGCCCG 300
DB 241 GAGAGAGCTTCACTCTCTCTGAGGAGGAGTGGGCGGCTGGGCTGTGCTCTGCCCG 300

QY 301 CAAAGGGAGAGCTGCCGCGACCCCAACCCCGGAGCTGTGCTGCGCATCCAGAGAGAGC 360
DB 301 CAAAGGGAGAGCTGCCGCGACCCCAACCCCGGAGCTGTGCTGCGCATCCAGAGAGAGC 360

QY 361 GAGCGCTGGGCGGAGCTGGGATGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
DB 361 GAGCGCTGGGCGGAGCTGGGATGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420

QY 421 CTGACACAGCTGACGCTCCAGCCCGCTGTCTACCGCAGGGGCGGCGGCGGCGGCGG 480
DB 421 CTGACACAGCTGACGCTCCAGCCCGCTGTCTACCGCAGGGGCGGCGGCGGCGGCGG 480

QY 481 GCGGCGGCGGCGGCGGCTTCAGAGAGGAGCGGCGGCTGTCTTCCAGGCGGCGGCGGCGG 540
DB 481 GCGGCGGCGGCGGCGGCTTCAGAGAGGAGCGGCGGCTGTCTTCCAGGCGGCGGCGGCGG 540

QY 541 CACAGGAGCTGAGGCTCACTTTGTGGCGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 600
DB 541 CACAGGAGCTGAGGCTCACTTTGTGGCGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 600

QY 601 CTAGAGCGGCTGCTCTGAGGCGGCTTCCCTCATTTGCCAGCTTCAGGAGGCTG 660
DB 601 CTAGAGCGGCTGCTCTGAGGCGGCTTCCCTCATTTGCCAGCTTCAGGAGGCTG 660

QY 661 GAGCAGCGCTGGAGCAACCCCGTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
DB 661 GAGCAGCGCTGGAGCAACCCCGTGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720


```
QY 1741 CAGAGGACCGCTGCTGGAGACGCTGCCGAGCTCACCCTTCCTGCTGGAGCTTTC 1800
Db 1741 CAGAGGACCGCTGCTGGAGACGCTGCCGAGCTCACCCTTCCTGCTGGAGCTTTC 1800
QY 1801 GGTACCTCGTGTTCCTAGTCACTACAAAGTGGCTGTGTCTGGGCTGCCAGGCGGCC 1860
Db 1801 GGTACCTCGTGTTCCTAGTCACTACAAAGTGGCTGTGTCTGGGCTGCCAGGCGGCC 1860
QY 1861 TCGCCAGCATTCCTATCCATCTTCAATCAATGTTCTCTTCTCCACAGCCCGCAGAAC 1920
Db 1861 TCGCCAGCATTCCTATCCATCTTCAATCAATGTTCTCTTCTCCACAGCCCGCAGAAC 1920
QY 1921 AGGTGCTCTACCCCGGAGGAGGTGGTCCAGGCCAGCTGTGGTCTTGGGCTTGGCC 1980
Db 1921 AGGTGCTCTACCCCGGAGGAGGTGGTCCAGGCCAGCTGTGGTCTTGGGCTTGGCC 1980
QY 1981 ATGTGCTCCATCTGCTGTGGCAGACCCCTGCACCTGTGTGACCCAGCCCGCGCCG 2040
Db 1981 ATGTGCTCCATCTGCTGTGGCAGACCCCTGCACCTGTGTGACCCAGCCCGCGCCG 2040
QY 2041 CTGCGAGGAGGCGCGCTGACCCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100
Db 2041 CTGCGAGGAGGCGCGCTGACCCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100
QY 2101 GAGCATCTGTGAATGCTGAGCTCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
Db 2101 GAGCATCTGTGAATGCTGAGCTCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
QY 2161 GAGGAGGCGGAGCTGCTGCTCCCTCGAGGTCTCATGACAGGAGGAGGAGGAGGAG 2220
Db 2161 GAGGAGGCGGAGCTGCTGCTCCCTCGAGGTCTCATGACAGGAGGAGGAGGAGGAG 2220
QY 2221 TTCTGCTGGGCTGCGTCTCAACACCGCTCTCTTCTGCGCTGTGGGCTTGGGCTG 2280
Db 2221 TTCTGCTGGGCTGCGTCTCAACACCGCTCTCTTCTGCGCTGTGGGCTTGGGCTG 2280
QY 2281 GCCCAGCCGAGCTGCTCGAGGTCTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340
Db 2281 GCCCAGCCGAGCTGCTCGAGGTCTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340
QY 2341 GGCCTGAGGAGTGGGCTGGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
Db 2341 GGCCTGAGGAGTGGGCTGGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
QY 2401 ATGACCGTGGCTATCTGCTGGTGTGATGAGGAGGAGTCTCAGCTTCTGTCAGCGCTG 2460
Db 2401 ATGACCGTGGCTATCTGCTGGTGTGATGAGGAGGAGTCTCAGCTTCTGTCAGCGCTG 2460
QY 2461 CTGCACTGGTGGAAATTCAGAAACAGTTTCTACTAGGAGGAGGAGGAGGAGGAGGAG 2520
Db 2461 CTGCACTGGTGGAAATTCAGAAACAGTTTCTACTAGGAGGAGGAGGAGGAGGAGGAG 2520
QY 2521 TTCACTTCTGCTGCCAGATGACTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2580
Db 2521 TTCACTTCTGCTGCCAGATGACTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2580
QY 2581 GACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
Db 2581 GACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
```

RESULT 9
US-11-122-329-72
; Sequence 72, Application US/11122329
; Publication No. US20060019272A1
; GENERAL INFORMATION:
; APPLICANT: Geraci, Mark
; APPLICANT: Bull, Todd
; APPLICANT: Voelkel, Norbert
; APPLICANT: Coldren, Chris
; TITLE OF INVENTION: Diagnosis of Disease and Monitoring of Therapy Using Gene
; TITLE OF INVENTION: Expression Analysis of Peripheral Blood Cells

FILE REFERENCE: 2848-54
; CURRENT APPLICATION NUMBER: US/11/122,329
; PRIOR FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/568,129
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 72
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-122-329-72

Query Match 100.0%; Score 2640; DB 16; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CGCGCTGCGCGGACGCGGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
Db 1 CGCGCTGCGCGGACGCGGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
QY 61 GGCTCCATGTTCCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
Db 61 GGCTCCATGTTCCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
QY 121 GCCTACACCTGCTGAGTGTGGGCGAGCTGGGCGCTGTGGAGTTTCAGAGACCTCAAC 180
Db 121 GCCTACACCTGCTGAGTGTGGGCGAGCTGGGCGCTGTGGAGTTTCAGAGACCTCAAC 180
QY 181 GCCTCGGTGAGCGCTTCCAGAGACGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Db 181 GCCTCGGTGAGCGCTTCCAGAGACGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
QY 241 GAGAAGACCTTCACTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
Db 241 GAGAAGACCTTCACTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
QY 301 CAAAGGAGGAGGCTGCGGCAACCCCAACCCCGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 301 CAAAGGAGGAGGCTGCGGCAACCCCAACCCCGGAGGAGGAGGAGGAGGAGGAGGAGG 360
QY 361 GAGCGCTGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
Db 361 GAGCGCTGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
QY 421 CTGACACAGCTGAGCTTCCAGCGCGCTGCTACGCGAGGAGGAGGAGGAGGAGGAGGAG 480
Db 421 CTGACACAGCTGAGCTTCCAGCGCGCTGCTACGCGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 481 GCGCGCCACACAGATGGGCGCTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
Db 481 GCGCGCCACACAGATGGGCGCTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
QY 541 CACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Db 541 CACCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 601 CTAGAGCGCTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
Db 601 CTAGAGCGCTGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
QY 661 GAGCAGCGCTGGAGCAACCCCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
Db 661 GAGCAGCGCTGGAGCAACCCCGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
QY 721 TCCCTACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Db 721 TCCCTACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
QY 781 CAGCTCTTCCCGTTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Db 781 CAGCTCTTCCCGTTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
```

```
QY 841 CAGCAGAGCAGAGCTGAGAGGTCTCGGGAGACAGAGCGGTCTTGAGCCAGGTG 900
Db 841 CAGCAGAGCAGAGCTGAGAGGTCTCGGGAGACAGAGCGGTCTTGAGCCAGGTG 900
QY 901 CTAGGCCGGGTGTGAGCTGTGCTGCCGACAGGCGAGGTGCAAGATGAAGGCC 960
Db 901 CTAGGCCGGGTGTGAGCTGTGCTGCCGACAGGCGAGGTGCAAGATGAAGGCC 960
QY 961 GTGTACTCTGGCCTGAACAGGTGACAGGTGACACCAACAGGTGCTCATTTGCCGAG 1020
Db 961 GTGTACTCTGGCCTGAACAGGTGACAGGTGACACCAACAGGTGCTCATTTGCCGAG 1020
QY 1021 GCCTGTGTCTGTGCGAGACTGCTGCCGCTGCGAGAGGCTCTGCGGACAGCTCGATG 1080
Db 1021 GCCTGTGTCTGTGCGAGACTGCTGCCGCTGCGAGAGGCTCTGCGGACAGCTCGATG 1080
QY 1081 GAGGAGGAGTGTGAGTGTGCTGATCCGCTGCTGCGGAGCATGCTGCGGACACTC 1140
Db 1081 GAGGAGGAGTGTGAGTGTGCTGATCCGCTGCTGCGGAGCATGCTGCGGACACTC 1140
QY 1141 ATCCGACCAACCGCTTACGCGCAGCTTCCAGGGCATGTCGATCGCTACGGCTGGGC 1200
Db 1141 ATCCGACCAACCGCTTACGCGCAGCTTCCAGGGCATGTCGATCGCTACGGCTGGGC 1200
QY 1201 CGCTACAGAGGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCTGTTTGT 1260
Db 1201 CGCTACAGAGGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCTGTTTGT 1260
QY 1261 GTGATGTTGCGGATGTGGGCAACGGCTGTGCTCATGTTCTCTTGGCTGCGCATGTC 1320
Db 1261 GTGATGTTGCGGATGTGGGCAACGGCTGTGCTCATGTTCTCTTGGCTGCGCATGTC 1320
QY 1321 CTTGCGGAGAACCGGCTGTGAAGCGCGGAGAGAGTCTGGCAGACTTCTTTC 1380
Db 1321 CTTGCGGAGAACCGGCTGTGAAGCGCGGAGAGAGTCTGGCAGACTTCTTTC 1380
QY 1381 AGGGGCGGTACTGCTGCTGCTTATGGGCTGTTCTCCATCTACACCGGCTTCTATC 1440
Db 1381 AGGGGCGGTACTGCTGCTGCTTATGGGCTGTTCTCCATCTACACCGGCTTCTATC 1440
QY 1441 AACGAGTGTTCAGTGTGCGGCAACAGCATCTTCCCTCGGGCTGAGTGTGCGCGCATG 1500
Db 1441 AACGAGTGTTCAGTGTGCGGCAACAGCATCTTCCCTCGGGCTGAGTGTGCGCGCATG 1500
QY 1501 GCCAACAGTCTGGCTGAGTGTGATGCTTCTGGGCGGACGACGATGCTTACCTGAT 1560
Db 1501 GCCAACAGTCTGGCTGAGTGTGATGCTTCTGGGCGGACGACGATGCTTACCTGAT 1560
QY 1561 CCCAACGTCAACCGGTGTCTCTGGGACCTTACCCCTTTGGCATCGATCTATTTGGAGC 1620
Db 1561 CCCAACGTCAACCGGTGTCTCTGGGACCTTACCCCTTTGGCATCGATCTATTTGGAGC 1620
QY 1621 CTGGCTGCCAACCATGAGTGTCTCAACTCTCTTCAAGTGAAGTGTGCGTCACTCTG 1680
Db 1621 CTGGCTGCCAACCATGAGTGTCTCAACTCTCTTCAAGTGAAGTGTGCGTCACTCTG 1680
QY 1681 GCGCTGTGCACATGCGCTTTGGGCTGTGCTCGAGTCTTCAACACAGTGCATTTGCG 1740
Db 1681 GCGCTGTGCACATGCGCTTTGGGCTGTGCTCGAGTCTTCAACACAGTGCATTTGCG 1740
QY 1741 CAGAGGACCGGCTGTGCTGAGAGCGCTGCGGAGCTCACCTTCTGCTGGGACTCTTTC 1800
Db 1741 CAGAGGACCGGCTGTGCTGAGAGCGCTGCGGAGCTCACCTTCTGCTGGGACTCTTTC 1800
QY 1801 GGTTACTCTGTTTCTAGTCACTACAGTGTGCTGTGCTGGCTGCGAGGGCGGC 1860
Db 1801 GGTTACTCTGTTTCTAGTCACTACAGTGTGCTGTGCTGGCTGCGAGGGCGGC 1860
QY 1861 TCGCCGAGCATCTCATCACTTCTCAACATGTTCTCTTCTCCACAGCCCAAGCAAC 1920
Db 1861 TCGCCGAGCATCTCATCACTTCTCAACATGTTCTCTTCTCCACAGCCCAAGCAAC 1920
QY 1921 AGGTGCTCTACCCCGGAGAGGTGGTCCAGGCGACGCTGCTGCTGCGCTTGGCC 1980
```

```
Db 1921 AGGTGCTCTACCCCGGAGAGGTGGTCCAGGCGACGCTGCTGCTGCGCTTGGCC 1980
QY 1981 ATGTGTGCCATCTCTGCTGTGGCAACCCCTGTGCAACCGCACCGCCCGCCG 2040
Db 1981 ATGTGTGCCATCTCTGCTGTGGCAACCCCTGTGCAACCGCACCGCCCGCCG 2040
QY 2041 CTGCGGAGGCGCGCTGACCGACAGAGGAGGAAACAGGCGGGTGTGCTGCACTGCT 2100
Db 2041 CTGCGGAGGCGCGCTGACCGACAGAGGAGGAAACAGGCGGGTGTGCTGCACTGCT 2100
QY 2101 GACGCATCTGTGAATGGCTGGAGCTCCATGAGGAGGAAAGGAGGCGCTGATGATGA 2160
Db 2101 GACGCATCTGTGAATGGCTGGAGCTCCATGAGGAGGAAAGGAGGCGCTGATGATGA 2160
QY 2161 GAGGAGGCGGAGCTGCTCCCTTCCAGGTGCTCATGCAACAGGCGCATCCACACATCGAG 2220
Db 2161 GAGGAGGCGGAGCTGCTCCCTTCCAGGTGCTCATGCAACAGGCGCATCCACACATCGAG 2220
QY 2221 TTCTGCTGGGCTGTGCTCCAAACCGCTCTACCTGCGCTGTGGGCGCTGAGCCCTG 2280
Db 2221 TTCTGCTGGGCTGTGCTCCAAACCGCTCTACCTGCGCTGTGGGCGCTGAGCCCTG 2280
QY 2281 GCCACGCGCCAGCTCTCCAGGTCTGTGGGCGCATGGTGATGCGATAGGCGCTGGGCGCTG 2340
Db 2281 GCCACGCGCCAGCTCTCCAGGTCTGTGGGCGCATGGTGATGCGATAGGCGCTGGGCGCTG 2340
QY 2341 GCGCGGAGGTGGGCGTGGCGCTGTGTGCTGTGCTCCCATCTTTGCGCGCTTTGCGCTG 2400
Db 2341 GCGCGGAGGTGGGCGTGGGCGTGTGTGCTGTGCTCCCATCTTTGCGCGCTTTGCGCTG 2400
QY 2401 ATGACCGTGGCTATCTGCTGTGATGAGGAGCTCTCAGCCTTCTGCAACCGCTGCGG 2460
Db 2401 ATGACCGTGGCTATCTGCTGTGATGAGGAGCTCTCAGCCTTCTGCAACCGCTGCGG 2460
QY 2461 CTGCACTCGGTGGAAATCCAGAACAAAGTTCTACTCAGGCGACGGGCTCAAGCTGATGCC 2520
Db 2461 CTGCACTCGGTGGAAATCCAGAACAAAGTTCTACTCAGGCGACGGGCTCAAGCTGATGCC 2520
QY 2521 TTCACTCTGCTGCCACAGATGATAGGGCCCATGCGAGTCTCTGCCAGACCTCTTCT 2580
Db 2521 TTCACTCTGCTGCCACAGATGATAGGGCCCATGCGAGTCTCTGCCAGACCTCTTCT 2580
QY 2581 GACCTCTCAGGCGAGAGGAGTAAGACGGTCCGCGCTGGCAAAAAA 2640
Db 2581 GACCTCTCAGGCGAGAGGAGTAAGACGGTCCGCGCTGGCAAAAAA 2640
```

RESULT 10

US-10-145-012-1
; Sequence 1, Application US/10145012
; Publication No. US20030124614A1
; GENERAL INFORMATION:
; APPLICANT: UTUKU et al.
; TITLE OF INVENTION: T-CELL MEMBRANE PROTEIN (TIRC7), PEPTIDES AND ANTIBODIES DERIVED
; FILE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 4400-0105P
; CURRENT APPLICATION NUMBER: US/10/145,012
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2488
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (537)..(2378)
US-10-145-012-1

Query Match 78.3%; Score 2068; DB 7; Length 2488;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 2105; Conservative 0; Mismatches 40; Indels 3; Gaps 1;

RESULT 11
US-11-126-866-1
; Sequence 1, Application US/111126866
; Publication No. US20050220789A1
; GENERAL INFORMATION:
; APPLICANT: UTUKU, et al., NALAN
; TITLE OF INVENTION: METHODS OF TREATMENT USING AN ANTIBODY THAT SPECIFICALLY BINDS TO
; FILE REFERENCE: 1472/71099-2A/JPW/AG
; CURRENT APPLICATION NUMBER: US/11/126,866
; CURRENT FILING DATE: 2005-05-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 2488
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (537)..(2378)
US-11-126-866-1

Query Match 78.3%; Score 2068; DB 13; Length 2488;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 2105; Conservative 0; Mismatches 40; Indels 3; Gaps 1;

| | | | |
|----|------|---|------|
| QY | 496 | GGGGCTCAGAGGAGCGCCCTGCTCCAGGCCCGGGGGCGCGCACAGGACCTGAGG | 555 |
| DB | 327 | GGGGCGCTGGTGGGGAGGACGGGAGGAGTGGAGAGCCCTGCCACGCCCGTGGCC | 386 |
| QY | 556 | GTCAACTTTGTGGCAGGTGCCGTGGAGCCCGCACAGGCCCTGCGCTAGAGCGCTGCTC | 615 |
| DB | 387 | GCCAGCTTTGTGCAGGTGCGGTGGAGCCCGCACAGGCCCTGCGCTAGAGCGCTGCTC | 446 |
| QY | 616 | TGGAGGCGCTGCGGGCTTCTCATTTGCCAGCTTCAGGAGCTGGAGCAGCGCTGGAG | 675 |
| DB | 447 | TGGAGGCGCTGCGGGCTTCTCATTTGCCAGCTTCAGGAGCTGGAGCAGCGCTGGAG | 506 |
| QY | 676 | CACCCGCTGACGGGAGCGCCAGCTGATGATGACCTTCTCATCTCTACTTGGGGTGAG | 735 |
| DB | 507 | CACCCGCTGACGGGAGCGCCAGCTGATGATGACCTTCTCATCTCTACTTGGGGTGAG | 566 |
| QY | 736 | CAGATCGGACAGAGATCGCGAAGATCACGGACTGCTTCCACTGCGCACGCTTCCCGTTT | 795 |
| DB | 567 | CAGATCGGACAGAGATCGCGAAGATCACGGACTGCTTCCACTGCGCACGCTTCCCGTTT | 626 |
| QY | 796 | CTGCAGCAGGAGGAGCGCCGCTCGGGGCGCTGCGAGAGCTGCAACAGCAGAGCAGGAG | 855 |
| DB | 627 | CTGCAGCAGGAGGAGCGCCGCTCGGGGCGCTGCGAGAGCTGCAACAGCAGAGCAGGAG | 686 |
| QY | 856 | CTGCAGGAGGTCTCGGGAGACAGAGCGGTTCTTGAGCCAGGTGCTAGGCGGGTGCTG | 915 |
| DB | 687 | CTGCAGGAGGTCTCGGGAGACAGAGCGGTTCTTGAGCCAGGTGCTAGGCGGGTGCTG | 746 |
| QY | 916 | CAGCTGCTGCCCGCAGGAGGTGAGGTGCACAGAGTGAAGGCGGTGTACCTTGGCCCTG | 975 |
| DB | 747 | CAGCTGCTGCCCGCAGGAGGTGAGGTGCACAGAGTGAAGGCGGTGTACCTTGGCCCTG | 806 |
| QY | 976 | AACAGGTGAGGTGAGCACACGCAAGTGCCTCATTTGCGAGGCTGTGTCTGTG | 1035 |
| DB | 807 | AACAGGTGAGGTGAGCACACGCAAGTGCCTCATTTGCGAGGCTGTGTCTGTG | 866 |
| QY | 1036 | CGAGACCTGCCCGCTGAGGAGCGCTGCGGACAGCTCGATGAGGAGGAGTGAGT | 1095 |
| DB | 867 | CGAGACCTGCCCGCTGAGGAGCGCTGCGGACAGCTCGATGAGGAGGAGTGAGT | 926 |
| QY | 1096 | GCGGTGGCTCACCGATCCCTGCGGGACATGCCCCCACCACACTCATCGCACCAACCGC | 1155 |
| DB | 927 | GCGGTGGCTCACCGATCCCTGCGGGACATGCCCCCACCACACTCATCGCACCAACCGC | 986 |
| QY | 1156 | TTACAGGCGAGCTTCAGGGCATCTGTGATTCGCTACGGGCTGGGCGCTACAGAGGTGTC | 1215 |
| DB | 987 | TTACAGGCGAGCTTCAGGGCATCTGTGATTCGCTACGGGCTGGGCGCTACAGAGGTGTC | 1046 |

| | | | |
|----|------|--|------|
| QY | 1216 | AAACCGCTCCCTACACCATCATCACCTTCCCTTCTGTTGCTGTGTGATGTTCCGGGAT | 1275 |
| DB | 1047 | AAACCGCTCCCTACACCATCATCACCTTCCCTTCTGTTGCTGTGTGATGTTCCGGGAT | 1106 |
| QY | 1276 | GTGGGCCACGGGCTCTCATGTTCTTTCGCTTGGCCATGGTCTTTCGGGAGAACCGA | 1335 |
| DB | 1107 | GTGGGCCACGGGCTCTCATGTTCTTTCGCTTGGCCATGGTCTTTCGGGAGAACCGA | 1166 |
| QY | 1336 | CGGGCTGTAAAGCGCGCAGAACAGATCTGGGAGACTTTCTTTCAGGGGCGCTACCTG | 1395 |
| DB | 1167 | CGGGCTGTAAAGCGCGCAGAACAGATCTGGGAGACTTTCTTTCAGGGGCGCTACCTG | 1226 |
| QY | 1396 | CTCTGCTTATGGGCTGTCTTCCATCTACACGGCTTCTCATCAACAGAGTCTTCACT | 1455 |
| DB | 1227 | CTCTGCTTATGGGCTGTCTTCCATCTACACGGCTTCTCATCAACAGAGTCTTCACT | 1286 |
| QY | 1456 | CGCGCCACAGCATCTTCCCTTCCGGCTGGAGTGTGGCGCCATGGCCAAACAGTCTGGC | 1515 |
| DB | 1287 | CGCGCCACAGCATCTTCCCTTCCGGCTGGAGTGTGGCGCCATGGCCAAACAGTCTGGC | 1346 |
| QY | 1516 | TGGAGTGTATGATTCCTTGGCCAGCACAGATCTTACCTTGGATCCCAACCTCACCGGT | 1575 |
| DB | 1347 | TGGAGTGTATGATTCCTTGGCCAGCACAGATCTTACCTTGGATCCCAACCTCACCGGT | 1406 |
| QY | 1576 | GTCTTCTTGGGACCTTACCCCTTGGCATCGATCTTATTTGGAGCTTGGCTGCCAACAC | 1635 |
| DB | 1407 | GTCTTCTTGGGACCTTACCCCTTGGCATCGATCTTATTTGGAGCTTGGCTGCCAACAC | 1466 |
| QY | 1636 | TTGAGCTTCTCAACTCTTCAAGATGAAGATGTCCGTCACTCTGGGGGTCTGTGCACATG | 1695 |
| DB | 1467 | TTGAGCTTCTCAACTCTTCAAGATGAAGATGTCCGTCACTCTGGGGGTCTGTGCACATG | 1526 |
| QY | 1696 | GCCTTTGGGGTGGTCTCGAGTCTTCAACCAAGTGCATTTTGGCCAGAGGACCGGGCTG | 1755 |
| DB | 1527 | GCCTTTGGGGTGGTCTCGAGTCTTCAACCAAGTGCATTTTGGCCAGAGGACCGGGCTG | 1586 |
| QY | 1756 | CTGCTGGAGAGCGCTGCCGAGCTACCTTCTGCTGGGACTTCTCGGTACCTCGTGTTC | 1815 |
| DB | 1587 | CTGCTGGAGAGCGCTGCCGAGCTACCTTCTGCTGGGACTTCTCGGTACCTCGTGTTC | 1646 |
| QY | 1816 | CTAGTCATCTACAAGTGGCTGTGTCTGGGTGCGCAGGGCGCGCTCG---CCCAGCATC | 1872 |
| DB | 1647 | CTAGTCATCTACAAGTGGCTGTGTCTGGGTGCGCAGGGCGCGCTCGGGCCCGAGCATC | 1706 |
| QY | 1873 | CTCATCCATCTCAATGTTCTCTTCTCCCAAGCAGCCCGAGCAACAGGTGTCTCTAC | 1932 |
| DB | 1707 | CTCATCCATCTCAATGTTCTCTTCTCCCAAGCAGCCCGAGCAACAGGTGTCTCTAC | 1766 |
| QY | 1933 | CCCCGCGAGGAGTGGTCCAGGCCACGCTGGTGTCTTGGCTTGGCCATGGTGGCCATC | 1992 |
| DB | 1767 | CCCCGCGAGGAGTGGTCCAGGCCACGCTGGTGTCTTGGCTTGGCCATGGTGGCCATC | 1826 |
| QY | 1993 | CTGCTGCTTGGCACACCCCTGTCACCTGTGCACCGCCACCGCGCGCGCTTGGCGAGG | 2052 |
| DB | 1827 | CTGCTGCTTGGCACACCCCTGTCACCTGTGCACCGCCACCGCGCGCGCTTGGCGAGG | 1886 |
| QY | 2053 | CCCGCTGACCGCAGGAGGAGGAAAAGAGCGCGGGTGTGTGAGCTTGCCTGACGATCTGTG | 2112 |
| DB | 1887 | CCCGCTGACCGCAGGAGGAGGAAAAGAGCGCGGGTGTGTGAGCTTGCCTGACGATCTGTG | 1946 |
| QY | 2113 | AATGCTCGAGTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG | 2172 |
| DB | 1947 | AATGCTCGAGTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG | 2006 |
| QY | 2173 | CTGCTCCCTCCGAGGTCTCATGCAACAGGCGCATCCACACCATCGAGTCTTCCCTGGGC | 2232 |
| DB | 2007 | CTGCTCCCTCCGAGGTCTCATGCAACAGGCGCATCCACACCATCGAGTCTTCCCTGGGC | 2066 |
| QY | 2233 | TGCTGTCTCAACACCGCTCTCTACTGCTGCTGTGGGCGCTTGGAGCTTGGCCACGCGC | 2292 |
| DB | 2067 | TGCTGTCTCAACACCGCTCTCTACTGCTGCTGTGGGCGCTTGGAGCTTGGCCACGCGC | 2126 |

| | | | |
|----|------|---|------|
| Qy | 2293 | CTGTCCGAGGTTCTGTGGCCCATGGTGTATGGCGATAGCCCTGGCCCTGGCCCGGGAGGTG | 2352 |
| Db | 2127 | CTGTCCGAGGTTCTGTGGCCCATGGTGTATGGCGATAGCCCTGGCCCTGGCCCGGGAGGTG | 2186 |
| Qy | 2353 | GGCGTGGCGGTGTGGTGCTGGTCCCATCTTTTCCGCGCTTTTGGCGGTGATGACCGTGGCT | 2412 |
| Db | 2187 | GGCGTGGCGGTGTGGTGCTGGTCCCATCTTTTCCGCGCTTTTGGCGGTGATGACCGTGGCT | 2246 |
| Qy | 2413 | ATCTGTCTGGTGATGGAGGACTCTGACGCTTCTGTCAGCCCTGCGGCTGCATCGGTG | 2472 |
| Db | 2247 | ATCTGTCTGGTGATGGAGGACTCTGACGCTTCTGTCAGCCCTGCGGCTGCATCGGTG | 2306 |
| Qy | 2473 | GAATTCCAGAA CAAGTTCTTACTCAGGCA CGGGCTACAAGCTGAGTCCCTTCACCTTCGCT | 2532 |
| Db | 2307 | GAATTCCAGAA CAAGTTCTTACTCAGGCA CGGGCTACAAGCTGAGTCCCTTCACCTTCGCT | 2366 |
| Qy | 2533 | GCCACAGATGACTAGGCGCCACTGCAGTCTCTGCGACGCTCCCTTCCTGACCTCTGAGGC | 2592 |
| Db | 2367 | GCCACAGATGACTAGGCGCCACTGCAGTCTCTGCGACGCTCCCTTCCTGACCTCTGAGGC | 2426 |
| Qy | 2593 | AGGAGAGGAATAAAGACGGTCCGCCCTGGCAAAAAAAAAAAAAAAAAAAAA | 2640 |
| Db | 2427 | AGGAGAGGAATAAAGACGGTCCGCCCTGGCAAAAAAAAAAAAAAAAAAAAA | 2474 |

RESULT 12

US-11-126-841A-1
 ; Sequence: 1, Application US/111126841A
 ; Publication No. US20050271659A1
 ; GENERAL INFORMATION:
 ; APPLICANT: UTKU, et al., NALAN
 ; TITLE OF INVENTION: METHODS OF TREATMENT USING AN ANTIBODY THAT SPECIFICALLY BINDS TO
 ; TITLE OF INVENTION: AN EXTRACELLULAR DOMAIN OF A TIRC7 MEMBRANE PROTEIN
 ; FILE REFERENCE: 1472/71099-ZB/jpw/AG
 ; CURRENT APPLICATION NUMBER: US/11/126,841A
 ; CURRENT FILING DATE: 2005-05-10
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 1
 ; LENGTH: 2488
 ; TYPE: DNA
 ; ORGANISM: human
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (537) .. (2378)
 US-11-126-841A-1

| | Query Match | 78.3% | Score 2068; | DB 15; | Length 2488; |
|----|-----------------------|---|----------------|-----------|--------------|
| | Best Local Similarity | 98.0%; | Pred. No. 0; | | |
| | Matches 2105; | Conservative 0; | Mismatches 40; | Indels 3; | Gaps 1; |
| Qy | 496 | GGGGCCCTCAGAGGACGCCCTCTCCAGGCCCGGGGGCGCACACAGGACCTGAGG | 555 | | |
| | | | | | |
| Db | 327 | GGGGGCCCTGTGTGGGGGAGGACGAGCGTGTGGAGCAGCCCTGCCCAGCCCGCTGGCC | 386 | | |
| | | | | | |
| Qy | 556 | GTCAACTTTTGGCAGGTGCGGTGGAGCCCAACAAGGCCCTGCGCTTAGAGCGCCTGCTC | 615 | | |
| | | | | | |
| Db | 387 | GCCAGCTTTTGTGGCAGGTGCGGTGGAGCCCAACAAGGCCCTGCGCTTAGAGCGCCTGCTC | 446 | | |
| | | | | | |
| Qy | 616 | TGAGGGCCCTGCCCGCGCTTTCCTATTGCCAGCTTTCAGGGAGCTTGAGAGCGCGCTGGAG | 675 | | |
| | | | | | |
| Db | 447 | TGAGGGCCCTGCCCGCGCTTTCCTATTGCCAGCTTTCAGGGAGCTTGAGAGCGCGCTGGAG | 506 | | |
| | | | | | |
| Qy | 676 | CACCCCTGACGGGCGAGCCAGCCACGTGATGACCTTCTCTATCTCTACTGGGGTGAG | 735 | | |
| | | | | | |
| Db | 507 | CACCCCTGACGGGCGAGCCAGCCACGTGATGACCTTCTCTATCTCTACTGGGGTGAG | 566 | | |
| | | | | | |
| Qy | 736 | CAGATCGGACAGAGATCCGCAAGATCACGGAGTGTCTTCCACTGCCACGCTTCCCGTTT | 795 | | |
| | | | | | |
| Db | 567 | CAGATCGGACAGAGATCCGCAAGATCACGGAGTGTCTTCCACTGCCACGCTTCCCGTTT | 626 | | |
| | | | | | |
| Qy | 796 | CTGCAGCAGGAGGAGGCCCGCGCTTGGGGCCCTGCAGCAGCTGCAACAGCAGACCCAGGAG | 855 | | |
| | | | | | |

| | | |
|-----|--|------|
| 627 | CTGCAGCAGGAGGAGGCCCCGCTCGGGGCCCTTGACGACGCTGCAACAGCAGACGCCAGGAG | 686 |
| Qy | CTGCAGGAGGTCTCTCGGGGAGACAGACGGGTTCCTGAGCCAGCTGTGTAGGCCGGGTGCTG | 915 |
| | CTGCAGGAGGTCTCTCGGGGAGACAGACGGGTTCCTGAGCCAGGTGTCTAGCCGGGTGCTG | 746 |
| | CAGCTGTCTGCCCGCAGGGCAGGTGCAGGTCCAAAGATGAAGGCCGTGTACCTTGGCCCTG | 806 |
| Qy | AACCAGTGCAGCGTGAGCACACGCAACAAGTGCCTCATTTGCCAGAGCCCTGTGCTCTGTG | 1035 |
| | AACCAGTGCAGCGTGAGCACACGCAACAAGTGCCTCATTTGCCAGAGCCCTGTGCTCTGTG | 866 |
| | CGAGACCTGCGCCCTGCAAGGAGGCCCTTCGGGGAACAGTTCGATGGAGGAGGAGTGAGT | 1095 |
| Db | CGAGACCTGCGCCCTGCAAGGAGGCCCTTCGGGGAACAGTTCGATGGAGGAGGAGTGAGT | 926 |
| | GCCGTGCTCAACCGCATCCCTTCGGGGACATGCCCCCCACACTCATTCGCAACAACCGC | 986 |
| | TTCAAGGCCAGCTTCCAGGGCATCTGTGGATGTCCTGCTGCTGTGATGTTTCGGGGAT | 1155 |
| Qy | TTCAAGGCCAGCTTCCAGGGCATCTGTGGATGTCCTGCTGCTGTGATGTTTCGGGGAT | 1215 |
| | AACCCCGCTCCCTACACCATCATCACTTCCCTCTCTGTTTCTGCTGTGATGTTTCGGGGAT | 1275 |
| | AACCCCGCTCCCTACACCATCATCACTTCCCTCTCTGTTTCTGCTGTGATGTTTCGGGGAT | 1106 |
| Qy | GTGGGCCACCGGGTGCTCATGTTCTTTCGGCCCTGGCCCATGCTCTTTCGGAGAACCGA | 1335 |
| | GTGGGCCACCGGGTGCTCATGTTCTTTCGGCCCTGGCCCATGCTCTTTCGGAGAACCGA | 1166 |
| | CCGGCTGTGAAGCCCGCGCAGACGAGATCTGGCAGACTTTCTTTCAGGGGCCGCTACCTG | 1395 |
| Db | CCGGCTGTGAAGCCCGCGCAGACGAGATCTGGCAGACTTTCTTTCAGGGGCCGCTACCTG | 1226 |
| | CTCTGCTTATGGGCGTGTCTCCATCTACACGGCTTTCATCTTACAAACAGTGTCTTTCAGT | 1455 |
| | CTCTGCTTATGGGCGTGTCTCCATCTACACGGCTTTCATCTTACAAACAGTGTCTTTCAGT | 1286 |
| Qy | CGCGCCACCAAGCATCTTCCCTCGGGCTGGAGTGTGGCCGCATGTCCTTTCAGGAGTGTGC | 1515 |
| | CGCGCCACCAAGCATCTTCCCTCGGGCTGGAGTGTGGCCGCATGTCCTTTCAGGAGTGTGC | 1346 |
| | TGAGTGATGCTATCTTGGCCGAGCACAAGTCTTACCTTACCTTGGATCCCAACGTCACCGGT | 1575 |
| Db | TGAGTGATGCTATCTTGGCCGAGCACAAGTCTTACCTTGGATCCCAACGTCACCGGT | 1406 |
| | GTCTTCTCGGGACCTTACCCCTTTTGGCATCGATCTTATTTGGAGCCTGTGCTGCCAACAC | 1635 |
| | GTCTTCTCGGGACCTTACCCCTTTTGGCATCGATCTTATTTGGAGCCTGTGCTGCCAACAC | 1466 |
| Qy | TTGAGCTTCTCAACTCTTTCAAGATGAAGATGTCCTGCTCATCTCTGGGCGCTCGTGACATG | 1695 |
| | TTGAGCTTCTCAACTCTTTCAAGATGAAGATGTCCTGCTCATCTCTGGGCGCTCGTGACATG | 1526 |
| | GCCTTTGGGGTGTCTCTGGAGTCTTCAACACAGTGCACTTTTGGCCAGAGGACCGGGCTG | 1755 |
| Db | GCCTTTGGGGTGTCTCTGGAGTCTTCAACACAGTGCACTTTTGGCCAGAGGACCGGGCTG | 1586 |
| | CTGCTGGAGACGTGGCGGAGCTCACCTTCTGCTGGGACTCTTTCGGTTCACCTCGTGTC | 1815 |
| | CTGCTGGAGACGTGGCGGAGCTCACCTTCTGCTGGGACTCTTTCGGTTCACCTCGTGTC | 1646 |
| Qy | CTAGTCACTTACAAGTGGCTGTGTGTCTGGGCTGCCAGGGCCGCTCG--CCAGCATC | 1872 |
| | CTAGTCACTTACAAGTGGCTGTGTGTCTGGGCTGCCAGGGCCGCTCG--CCAGCATC | 1706 |
| | CTCATCCACTTCAACAATGTTCTTCTTCTCCACAGCCCCAGCAACAGGTGTCTTAC | 1932 |
| Db | CTCATCCACTTCAACAATGTTCTTCTTCTCCACAGCCCCAGCAACAGGTGTCTTAC | 1766 |

```
QY 1933 CCCCGCAGAGTGTGTCAGGCGACGCTGCTGCTCTGCGCCTTGCCATGTTGCCATC 1992
Db 1767 CCCCGCAGAGTGTGTCAGGCGACGCTGCTGCTGCTCTGCGCCTTGCCATGTTGCCATC 1826
QY 1993 CTGCTCTTTGGCACAACCCCTGCACCTGCTGCACCGCCACCGCGCCCTGCGGAGGAGG 2052
Db 1827 CTGCTCTTTGGCACAACCCCTGCACCTGCTGCACCGCCACCGCGCCCTGCGGAGGAGG 1886
QY 2053 CCGCTGACCGACAGAGGAAACAAAGGCGGGTGTCTGGACCTGCTGACGATCTGTG 2112
Db 1887 CCGCTGACCGACAGAGGAAACAAAGGCGGGTGTCTGGACCTGCTGACGATCTGTG 1946
QY 2113 AATGGCTGAGCTCCGATCAGAGAAAGGAGGGGCTCGATGATGAAGAGGAGCGCGAG 2172
Db 1947 AATGGCTGAGCTCCGATCAGAGAAAGGAGGGGCTCGATGATGAAGAGGAGCGCGAG 2006
QY 2173 CTGCTCCCTCCGAGGTGCTCATGACACAGGGCATCCACACCATCGAGTTCTGCTGGGC 2232
Db 2007 CTGCTCCCTCCGAGGTGCTCATGACACAGGGCATCCACACCATCGAGTTCTGCTGGGC 2066
QY 2233 TGGGTCTCAACACCGCTCTACCTGCGCTGTGGGCGCTGAGCCTTGCGCCACGCGCCAG 2292
Db 2067 TGGGTCTCAACACCGCTCTACCTGCGCTGTGGGCGCTGAGCCTTGCGCCACGCGCCAG 2126
QY 2293 CTGCTCGAGGTTCTGTGGGCCATGTTGATGCGCATAGGCTGCGCTGGCGCGGAGGTG 2352
Db 2127 CTGCTCGAGGTTCTGTGGGCCATGTTGATGCGCATAGGCTGCGCTGGCGCGGAGGTG 2186
QY 2353 GCGGTGGGCGTGTGTTGCTGCCATCTTTTGGCGCTTTTGGCGTATGACCGTGGCT 2412
Db 2187 GCGGTGGGCGTGTGTTGCTGCCATCTTTTGGCGCTTTTGGCGTATGACCGTGGCT 2246
QY 2413 ATCTGCTGGTGAATGAGGAGTCTCAGGCTTCTGCGACCGCTTGGCGTGCATCTGGGTG 2472
Db 2247 ATCTGCTGGTGAATGAGGAGTCTCAGGCTTCTGCGACCGCTTGGCGTGCATCTGGGTG 2306
QY 2473 GAATTCACAGAACAGTTCTACTCAGGACGGGCTACAGCTGAGTCCCTTACCTTCGCT 2532
Db 2307 GAATTCACAGAACAGTTCTACTCAGGACGGGCTACAGCTGAGTCCCTTACCTTCGCT 2366
QY 2533 GCCACAGATGACTAGGGCCACTGAGGTCTCTGCCAGACCTCTTCTGACCTCTGAGGC 2592
Db 2367 GCCACAGATGACTAGGGCCACTGAGGTCTCTGCCAGACCTCTTCTGACCTCTGAGGC 2426
QY 2593 AGGAGGATGAAGACGTTCCGCTTGGCAAAAAA 2640
Db 2427 AGGAGGATGAAGACGTTCCGCTTGGCAAAAAA 2474

RESULT 13
US-10-145-012-12
; Sequence 12, Application US/10145012
; Publication No. US20030124614A1
; GENERAL INFORMATION:
; APPLICANT: UTU et al.
; TITLE OF INVENTION: T-CELL MEMBRANE PROTEIN (TIRC7), PEPTIDES AND ANTIBODIES DERIVED
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 4400-0105P
; CURRENT APPLICATION NUMBER: US/10/145,012
; CURRENT FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 2488
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (537)..(2378)
US-10-145-012-12
```

Query Match 78.3%; Score 2066.4; DB 7; Length 2488;

```
Best Local Similarity 98.0%; Pred. No. 0;
Matches 2104; Conservative 0; Mismatches 41; Indels 3; Gaps 1;
QY 496 GGGGCTCTAGAGAGACGCCCTGTCTCCAGGCCCCCGGGGGCCGACACAGGACCTTGAGG 555
Db 327 GGGGCTCTGGTGGGGAGGAGGAGGAGGAGGTTGGAGCAGCCCTTGCCAGCCCGTGGCC 386
QY 556 GTCAACTTTTGTGGCAGGTGCGGTGAGGCCCAAGAGGCCCTTGCCCTAGAGGCCCTGCTC 615
Db 387 GCCAGCTTTGTGGCAGGTGCGGTGAGGCCCAAGAGGCCCTTGCCCTAGAGGCCCTGCTC 446
QY 616 TGGAGGCTGCGCGGCTTCTCATTTGACGCTTTCAGGAGCTTGAGCAGCCGCTGGAG 675
Db 447 TGGAGGCTGCGCGGCTTCTCATTTGACGCTTTCAGGAGCTTGAGCAGCCGCTGGAG 506
QY 676 CACCCCTGACCGGCGAGCCACGCTGGATGACCTTCTCATCTCTTACTTCCGGGTGAG 735
Db 507 CACCCCTGACCGGCGAGCCAGCCACGTTGATGACCTTCTCATCTCTTACTTCCGGGTGAG 566
QY 736 CAGATCGGACAGAGATCGGCAAGATCAGGACTGCTTCCACTGCACTGCACTTCCGGTTT 795
Db 567 CAGATCGGACAGAGATCGGCAAGATCAGGACTGCTTCCACTGCACTGCACTTCCGGTTT 626
QY 796 CTGACGAGGCTCTCGGGAGACAGAGCGGTTCTTGAGCCAGGTGCTTAGCCGGGTGCTG 915
Db 687 CTGACGAGGCTCTCGGGAGACAGAGCGGTTCTTGAGCCAGGTGCTTAGCCGGGTGCTG 746
QY 916 CAGCTGCTGCGGCGAGGCGAGGTGCAAGTCCACAGATGAGGCGGTGTTACCTGCGCCTG 975
Db 747 CAGCTGCTGCGGCGAGGCGAGGTGCAAGTCCACAGATGAGGCGGTGTTACCTGCGCCTG 806
QY 976 AACAGTGCAGCGTGAGCACCAAGTGCCTCATTTGCCAGAGCCCTTGGTGTCTGTG 1035
Db 807 AACAGTGCAGCGTGAGCACCAAGTGCCTCATTTGCCAGAGCCCTTGGTGTCTGTG 866
QY 1036 CGAGACTGCGGCGCTGAGAGGCCCTTGAGGACAGCTCGATGAGAGGAGGTGAGT 1095
Db 867 CGAGACTGCGGCGCTGAGAGGCCCTTGAGGACAGCTCGATGAGAGGAGGTGAGT 926
QY 1096 GCGGTGGCTCACCGCATCCCTGCGGAGCATGCCCGCCACACTCATCCGACACCAACCG 1155
Db 927 GCGGTGGCTCACCGCATCCCTGCGGAGCATGCCCGCCACACTCATCCGACACCAACCG 986
QY 1156 TTACGCGCAGCTTCCAGGCGATCGTGGATCGCTACGCGGTGGGCGCTTACCAGGAGTTC 1215
Db 987 TTACGCGCAGCTTCCAGGCGATCGTGGATCGCTACGCGGTGGGCGCTTACCAGGAGTTC 1046
QY 1216 AACCCGCTTCCATACCATCATACCTTCCCTTCCCTTCTGTTGCTGTGATGTTCCGGGAT 1275
Db 1047 AACCCGCTTCCATACCATCATACCTTCCCTTCCCTTCTGTTGCTGTGATGTTCCGGGAT 1106
QY 1276 GTGGCCACCGGCTCATCTTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC 1335
Db 1107 GTGGCCACCGGCTCATCTTCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC 1166
QY 1336 CCGGTGTGAAGCGCGCAGAACAGATCTGGGACAGCTTTCTTCCAGGGGCGCTACCTG 1395
Db 1167 CCGGTGTGAAGCGCGCAGAACAGATCTGGGACAGCTTTCTTCCAGGGGCGCTACCTG 1226
QY 1396 CTCTGCTTATGGGCTGTTCTCATCTTACAGCGGCTTCTCATCAACAGAGTGTCTCAGT 1455
Db 1227 CTCTGCTTATGGGCTGTTCTCTCATCTTACAGCGGCTTCTCATCAACAGAGTGTCTCAGT 1286
QY 1456 CGCGCCACAGCATCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC 1515
Db 1287 CGCGCCACAGCATCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC 1346
QY 1516 TGGAGTGTGATCTTCTTCCGCGCAGCACAGTGTCTTACCTTGGATCCCAACGTCACCGT 1575
```

```
Db 1347 TGGAGTGATGATTCCTGGGCCAGCACAGATGCTTACCTCGGATCCCAACGTACCGGT 1406
Qy 1576 GTCTTCTGGAGCCCTACCCCTTTGGCATCGATCCTATTTGGAGCCTGGCTGCCAACAC 1635
Db 1407 GTCTTCTGGAGCCCTACCCCTTTGGCATCGATCCTATTTGGAGCCTGGCTGCCAACAC 1466
Qy 1636 TTGAGCTTCTCAACTCCTTCAAGATGAAGATGTCGGTCATCTTGGGGGTGTCACATG 1695
Db 1467 TTGAGCTTCTCAACTCCTTCAAGATGAAGATGTCGGTCATCTTGGGGGTGTCACATG 1526
Qy 1696 GCCTTTGGGGTGTCTCGAGTCTTCAACACAGTCGACATTTGGCCAGAGGCACCGGTG 1755
Db 1527 GCCTTTGGGGTGTCTCGAGTCTTCAACACAGTCGACATTTGGCCAGAGGCACCGGTG 1586
Qy 1756 CTGCTGAGAGCTGCCGGAGCTCACCTTCTGCTGGGACTCTTTCGGTTACTCTGTTTC 1815
Db 1587 CTGCTGAGAGCTGCCGGAGCTCACCTTCTGCTGGGACTCTTTCGGTTACTCTGTTTC 1646
Qy 1816 CTAGTCATCTACAGTGGCTGTGTCTGGGCTGCCAGGGCGGCTCG---CCAGCATC 1872
Db 1647 CTAGTCATCTACAGTGGCTGTGTCTGGGCTGCCAGGGCGGCTCGGGCCCCAGATC 1706
Qy 1873 CTCATCACTTTCATCAACATGTTCTCTTCTTCCACAGCCCCACCAACAGGCTGCTCTAC 1932
Db 1707 CTCATCACTTTCATCAACATGTTCTCTTCTTCCACAGCCCCACCAACAGGCTGCTCTAC 1766
Qy 1933 CCCCAGAGAGTGTCTCAGGCCACAGCTGGTGTCTTGGCTTGGCCATGTTGCCCATC 1992
Db 1767 CCCCAGAGAGTGTCTCAGGCCACAGCTGGTGTCTTGGCTTGGCCATGTTGCCCATC 1826
Qy 1993 CTGCTGCTTGGCAGACCCCTGACCTCTGTCACCGCCACCGCCCGGCTTGGAGAGG 2052
Db 1827 CTGCTGCTTGGCAGACCCCTGACCTCTGTCACCGCCACCGCCCGGCTTGGAGAGG 1886
Qy 2053 CCGCTGACCGACAGGAGGAAAAAAGCCCGGGTGTCTGACCTGCTGACGATCTGTG 2112
Db 1887 CCGCTGACCGACAGGAGGAAAAAAGCCCGGGTGTCTGACCTGCTGACGATCTGTG 1946
Qy 2113 AATGGCTGAGCTCCGATGAGGAAAAAGCCAGGGGCTCGATGATGAAGAGGAGCCGAG 2172
Db 1947 AATGGCTGAGCTCCGATGAGGAAAAAGCCAGGGGCTCGATGATGAAGAGGAGCCGAG 2006
Qy 2173 CTGCTCCCTCCGAGGTCTCATGACAGGCCATCCACACCATCGATGTTTGGCTGGG 2232
Db 2007 CTGCTCCCTCCGAGGTCTCATGACAGGCCATCCACACCATCGATGTTTGGCTGGG 2066
Qy 2233 TGCGTCTCCAAACACCGCTCTCTACCTGCGCTGTGGGCCCTGAGCCTTGGCCCAAGCCAG 2292
Db 2067 TGCGTCTCCAAACACCGCTCTCTACCTGCGCTGTGGGCCCTGAGCCTTGGCCCAAGCCAG 2126
Qy 2293 CTCTCGAGGTTCTGTGGCCATGTGTATGCGCATAGGCCCTGGGCCCTGGGCCGGAGGTG 2352
Db 2127 CTCTCGAGGTTCTGTGGCCATGTGTATGCGCATAGGCCCTGGGCCCTGGGCCGGAGGTG 2186
Qy 2353 GCGGTGGCGCTGTGGTGTGGTCCCATCTTTGCCGCTTTCGCGTGATGACCGTGGCT 2412
Db 2187 GCGGTGGCGCTGTGGTGTGGTCCCATCTTTGCCGCTTTCGCGTGATGACCGTGGCT 2246
Qy 2413 ATCTCTGTGTATGAGGAGTCTCAGCCTTCTGACGCTTCTGACGCTTGGGCTGCACTGGGTG 2472
Db 2247 ATCTCTGTGTATGAGGAGTCTCAGCCTTCTGACGCTTGGGCTGCACTGGGTG 2306
Qy 2473 GAATTCAGAACTTCTACTCAGGACAGGGCTTACAGCTGAGTCCCTTTACCTTCGCT 2532
Db 2307 GAATTCAGAACTTCTACTCAGGACAGGGCTTACAGCTGAGTCCCTTTACCTTCGCT 2366
Qy 2533 GCCACAGATGACTAGGGCCACTGACAGTCTTCCAGACCTCTTCTGACCTCTGAGGC 2592
Db 2367 GCCACAGATGACTAGGGCCACTGACAGTCTTCCAGAGCTCTTCTGACCTCTGAGGC 2426
Qy 2593 AGGAGAGGAATAAAGACGGTCCGCTTGGCAAAAAAAGGAAAAA 2640
Db 2427 AGGAGAGGAATAAAGACGGTCCGCTTGGCAAAAAAAGGAAAAA 2474
```

RESULT 14

```
US-11-126-866-12
; Sequence 12, Application US/111126866
; Publication No. US20050220789A1
; GENERAL INFORMATION:
; APPLICANT: UTOKU, et al., NALAN
; TITLE OF INVENTION: METHODS OF TREATMENT USING AN ANTIBODY THAT SPECIFICALLY BINDS
; TITLE OF INVENTION: AN EXTRACELLULAR DOMAIN OF A TIRC7 MEMBRANE PROTEIN
; FILE REFERENCE: 1472/71099-2A/JPW/AG
; CURRENT APPLICATION NUMBER: US/11/126,866
; CURRENT FILING DATE: 2005-05-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12
; LENGTH: 2488
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic oligonucleotide
; NAME/KEY: CDS
; LOCATION: (537)..(2378)
US-11-126-866-12
```

```
Query Match 78.3%; Score 2066.4; DB 13; Length 2488;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 2104; Conservative 0; Mismatches 41; Indels 3; Gaps 1;

Qy 496 GGGGCTTCAGAGAGGACGCCCCCTGCTCCAGGCCCCGGGGGGCGCACACAGGACCTGAGG 555
Db 327 GGGGGGCTGTGTGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 386
Qy 556 GTCAACTTTTGTGCGAGTGCCTGGAGCCCAAGAGCCCTGCCCTAGAGCGCTGCTC 615
Db 387 GCCAGCTTTTGTGGAGGTGCTGGAGCCCAAGAGCCCTGCCCTAGAGCGCTGCTC 446
Qy 616 TGGAGGCGCTGCGCGCGCTTCTCTATTGCGAGCTTCAGGGAGCTTGGAGAGCGGCTGGAG 675
Db 447 TGGAGGCGCTGCGCGCGCTTCTCTATTGCGAGCTTCAGGGAGCTTGGAGAGCGGCTGGAG 506
Qy 676 CACCCGTGAGCGGGGAGAGCCAGCGATGAGTGAATGACCTTCTCATCTCTACTGCGGTGAG 735
Db 507 CACCCGTGAGCGGGGAGAGCCAGCGATGAGTGAATGACCTTCTCATCTCTACTGCGGTGAG 566
Qy 736 CAGATCGGACAGAGATCCGACAGATCAGGAGTGTCTTCCACTGCCACGCTTTCGCCGTTT 795
Db 567 CAGATCGGACAGAGATCCGACAGATCAGGAGTGTCTTCCACTGCCACGCTTTCGCCGTTT 626
Qy 796 CTGACAGAGAGAGGAGCGCCCTCGGGGCGCTGCAGCAGCTGCAACAGCAGAGAGCCAGAG 855
Db 627 CTGACAGAGAGAGGAGCGCCCTCGGGGCGCTGCAGCAGCTGCAACAGCAGAGAGCCAGAG 686
Qy 856 CTGACAGAGTCTCTCGGGAGACAGAGCGGTTCTTGAAGCAGGTGCTTGAAGCGGGTGTG 915
Db 687 CTGACAGAGTCTCTCGGGAGACAGAGCGGTTCTTGAAGCAGGTGCTTGAAGCGGGTGTG 746
Qy 916 CAGCTCTCGCGCCAGGCGAGGTGAGGTCCAGAGTGAAGGCGCTGTACTTGGGCCCTG 975
Db 747 CAGCTCTCGCGCCAGGCGAGGTGAGGTCCAGAGTGAAGGCGCTGTACTTGGGCCCTG 806
Qy 976 AACAGTGCAGCTGAGCAGCAGCAGAGTGTCTCATTTGCCAGGCGCTGGTGTCTGTG 1035
Db 807 AACAGTGCAGCTGAGCAGCAGCAGAGTGTCTCATTTGCCAGGCGCTGGTGTCTGTG 866
Qy 1036 CGAGACTGCGGCCCTTGAAGGCGCTTGAAGGCGCTTGAAGGCGCTTGAAGGCGCTTGAAG 1095
Db 867 CGAGACTGCGGCCCTTGAAGGCGCTTGAAGGCGCTTGAAGGCGCTTGAAGGCGCTTGAAG 926
Qy 1096 GCGGTGGCTCACCGCATCCCTTGGCGGAGCATGCCGCCACACTCATTCGCCACCAACCGC 1155
Db 927 GCGGTGGCTCACCGCATCCCTTGGCGGAGCATGCCGCCACACTCATTCGCCACCAACCGC 986
```


QY 736 CAGATCGGACAGAGATCCGCAAGATCACCGACTGCTTCACTGCCAGCTCTCCCGTTT 795
Db 567 CAGATCGGACAGAGATCCGCAAGATCACCGACTGCTTCACTGCCAGCTCTCCCGTTT 626
QY 796 CTGACGACAGAGAGGCGCGCTCGGGCCCTGACGAGCTGCAACAGCAGAGCCAGGAG 855
Db 627 CTGACGACAGAGAGGCGCGCTCGGGCCCTGACGAGCTGCAACAGCAGAGCCAGGAG 686
QY 856 CTGACGAGAGTCTCGGGGAGACAGAGCGGTTCTGAGCCAGAGTGTAGGCCGGGTCTG 915
Db 687 CTGACGAGAGTCTCGGGGAGACAGAGCGGTTCTGAGCCAGAGTGTAGGCCGGGTCTG 746
QY 916 CAGCTGCTCGCCAGCGGAGGTGACGTCACAGATGAAGCGGTGTACCTGGCCCTG 975
Db 747 CAGCTGCTCGCCAGCGGAGGTGACGTCACAGATGAAGCGGTGTACCTGGCCCTG 806
QY 976 AACCAAGTGCAGCTGACGACACACCAAGTGCCTCATTTGCCGAGGCTGTGTCTGTG 1035
Db 807 AACCAAGTGCAGCTGACGACACACCAAGTGCCTCATTTGCCGAGGCTGTGTCTGTG 866
QY 1036 CGAGACCTGCGCCCTGCGAGGAGCCCTCGGGACAGCTCGATFGAGGAGGAGTGAAT 1095
Db 867 CGAGACCTGCGCCCTGCGAGGAGCCCTGCGAGACAGCTCGATFGAGGAGGAGTGAAT 926
QY 1096 GCCGTGCTCACCGCATCCCTGCGGACATGCCCCCAGCTCATCTCCGACCAACCCG 1155
Db 927 GCCGTGCTCACCGCATCCCTGCGGACATGCCCCCAGCTCATCTCCGACCAACCCG 986
QY 1156 TTACGCGCCAGCTTTCAGAGGATCGTGGATCGCTACGGCGTGGGCGCTACACAGAGGTC 1215
Db 987 TTACGCGCCAGCTTTCAGAGGATCGTGGATCGCTACGGCGTGGGCGCTACACAGAGGTC 1046
QY 1216 AACCCGCTCCCTACACCATCATCCTTCCCTTCTGTTGCTGTGATGTTTCGGGGAT 1275
Db 1047 AACCCGCTCCCTACACCATCATCCTTCCCTTCTGTTGCTGTGATGTTTCGGGGAT 1106
QY 1276 GTGGCCACCGGCTGCTCATGTTCTTTCGCGCTGCGCATGTCTTGGGAGAACCGA 1335
Db 1107 GTGGCCACCGGCTGCTCATGTTCTTTCGCGCTGCGCATGTCTTGGGAGAACCGA 1166
QY 1336 CCGGCTGTAAAGCCGCGAGAGAGATCTGGCAGACTTTCCTTTCAGGGGCGCTACCTG 1395
Db 1167 CCGGCTGTAAAGCCGCGAGAGAGATCTGGCAGACTTTCCTTTCAGGGGCGCTACCTG 1226
QY 1396 CTCCTGCTTAAGGCTGTTCTCCATCTACACCGGCTTCACTACACAGAGTGTCTCAGT 1455
Db 1227 CTCCTGCTTAAGGCTGTTCTCCATCTACACCGGCTTCACTACACAGAGTGTCTCAGT 1286
QY 1456 CGCGCCACAGCATCTTCCCTCGGGCTGAGGTGTGCGCCCATGGCCAAACCATGCTGGC 1515
Db 1287 CGCGCCACAGCATCTTCCCTCGGGCTGAGGTGTGCGCCCATGGCCAAACCATGCTGGC 1346
QY 1516 TGGAGTGATGATTCCTTGGGCCAGCAGCATGCTTACCTTGGATCCCAACGTCACCGGT 1575
Db 1347 TGGAGTGATGATTCCTTGGGCCAGCAGCATGCTTACCTTGGATCCCAACGTCACCGGT 1406
QY 1576 GTCTTCTGGGACCTTACCTTGGCATCGATCTTATTTGGAGCTTGGCTGCCAACAC 1635
Db 1407 GTCTTCTGGGACCTTACCTTGGCATCGATCTTATTTGGAGCTTGGCTGCCAACAC 1466
QY 1636 TTGAGCTTCTCAACTCTTCAAGATGAAGATGTCCGTCACTCTGGGCTGTGCAATG 1695
Db 1467 TTGAGCTTCTCAACTCTTCAAGATGAAGATGTCCGTCACTCTGGGCTGTGCAATG 1526
QY 1696 GCCTTTGGGCTGTCTCGGAGTCTTCAACAGCTGCACTTTGGCAGAGGACCCGGCTG 1755
Db 1527 GCCTTTGGGCTGTCTCGGAGTCTTCAACAGCTGCACTTTGGCAGAGGACCCGGCTG 1586
QY 1756 CTGCTGGAGACGCTGCGGAGCTCACCTTCTGTGGGACTCTTTCGGTTACCTCGTGTTC 1815
Db 1587 CTGCTGGAGACGCTGCGGAGCTCACCTTCTGTGGGACTCTTTCGGTTACCTCGTGTTC 1646
QY 1816 CTAGTCATCTAAGTGGCTGTGTCTGGGCTGCGAGGGCCGCGCTCG---CCAGCATC 1872

Db 1647 CTAGTCATCTAAGTGGCTGTGTCTGGGCTGCCAGGGCCGCTCTGGCCCCAGCATC 1706
QY 1873 CTATTCACATCTTCAATCAACATGTTCTCTTCCACAGCCCGCAGCAACAGCTCTCTAC 1932
Db 1707 CTATTCACATCTTCAATCAACATGTTCTCTTCCACAGCCCGCAGCAACAGCTCTCTAC 1766
QY 1933 CCCCGCAGAGGTGTCTCAGGCCACGCTCGTGGTCTCTGGCTTTGGCCATTTGGTCCCAC 1992
Db 1767 CCCCGCAGAGGTGTCTCAGGCCACGCTCGTGGTCTCTGGCTTTGGCCATTTGGTCCCAC 1856
QY 1993 CTGCTGCTTGGCACACCCCTGCACTGCTGCAACGGCCACCGCCGCGCTCTCGGAGAGG 2052
Db 1827 CTGCTGCTTGGCACACCCCTGCACTGCTGCAACGGCCACCGCCGCGCTCTCGGAGAGG 1886
QY 2053 CCGCTGTACCGACAGAGGAGGAAAAAAGGGCCGGGTGTCTGACCTGCTGACGATCTGTG 2112
Db 1887 CCGCTGTACCGACAGAGGAGGAAAAAAGGGCCGGGTGTCTGACCTGCTGACGATCTGTG 1946
QY 2113 AATGGCTGGAGTCCGATGAGGAAAAAGGCAAGGGGCTGTGATGATGAAGAGGAGGCCAG 2172
Db 1947 AATGGCTGGAGTCCGATGAGGAAAAAGGCAAGGGGCTGTGATGATGAAGAGGAGGCCAG 2006
QY 2173 CTGCTCCCTTCCGAGGTGCTCATGCAACAGGCCATCCACACCATCGAGTTCTGCTTGGC 2232
Db 2007 CTGCTCCCTTCCGAGGTGCTCATGCAACAGGCCATCCACACCATCGAGTTCTGCTTGGC 2066
QY 2233 TGCGTCTCCAAACACCGCTCTACTCTGCGCTGTGGGCCCTTGAGCCCTGGGCCACGCCAG 2292
Db 2067 TGCGTCTCCAAACACCGCTCTACTCTGCGCTGTGGGCCCTTGAGCCCTGGGCCACGCCAG 2126
QY 2293 CTGCTCCAGAGTCTGTGGGCCATGTTGATGCGCATAGGCCCTGGGCCCTGGGCCAGGAGTG 2352
Db 2127 CTGCTCCAGAGTCTGTGGGCCATGTTGATGCGCATAGGCCCTGGGCCCTGGGCCAGGAGTG 2186
QY 2353 GCGCTGGCGGCTGTGTGCTGGTCCCATCTTTGCGGCCCTTTGCGCTGATGACCGTGGCT 2412
Db 2187 GCGCTGGCGGCTGTGTGCTGGTCCCATCTTTGCGGCCCTTTGCGCTGATGACCGTGGCT 2246
QY 2413 ATCTGCTGTGATGAGGAGCTCTCAGCTCTTCTGACGCCCTGGGCCCTGCACTGGGTG 2472
Db 2247 ATCTGCTGTGATGAGGAGCTCTCAGCTCTTCTGACGCCCTGGGCCCTGCACTGGGTG 2306
QY 2473 GAATTCAGAACAAAGTTCTACTCAGGCAAGGCTCAAGCTGAGTCCCTTTCACCTTCGCT 2532
Db 2307 GAATTCAGAACAAAGTTCTACTCAGGCAAGGCTCAAGCTGAGTCCCTTTCACCTTCGCT 2366
QY 2533 GCCACAGATGACTAGGGCCCACTGCAAGTCTCTGCGAGACTCTCTTCTGACCTCTGAGGC 2592
Db 2367 GCCACAGATGACTAGGGCCCACTGCAAGTCTCTGCGAGACTCTCTTCTGACCTCTGAGGC 2426
QY 2593 AGGAGAGGAATAAAGACGGTCCGCTTGGCAAAAAAAGGAAAAA 2640
Db 2427 AGGAGAGGAATAAAGACGGTCCGCTTGGCAAAAAAAGGAAAAA 2474

Search completed: June 30, 2006, 06:17:08
Job time : 3197 secs

THIS PAGE BLANK (USPTO)

QY 241 GAGAGACCTTACCTTCTGACGAGGAGGTGCGCGGCTGCGGTCTGCTGCCCGC 300
DB 241 GAGAGACCTTACCTTCTGACGAGGAGGTGCGCGGCTGCGGTCTGCTGCCCGC 300
QY 301 CCAAGGGAGGCTGCGGACCCCAACCCCGGACCTGCTGCGCATCCAGGAGGAGCG 360
DB 301 CCAAGGGAGGCTGCGGACCCCAACCCCGGACCTGCTGCGCATCCAGGAGGAGCG 360
QY 361 GAGCGCTGGCCAGAGCTGCGGGATGTCGGGGCAACAGCAGGCGCTTGGGGCCAG 420
DB 361 GAGCGCTGGCCAGAGCTGCGGGATGTCGGGGCAACAGCAGGCGCTTGGGGCCAG 420
QY 421 CTGACACAGCTGACGCTCACGCGCGCTGCTACGCGAGGCGCATGAACCTCAGCTGGCA 480
DB 421 CTGACACAGCTGACGCTCACGCGCGCTGCTACGCGAGGCGCATGAACCTCAGCTGGCA 480
QY 481 GCGGCCACACAGATGGGCTCTAGAGAGGACGCCCTGCTCCAGGCCCCCGGGGGCGG 540
DB 481 GCGGCCACACAGATGGGCTCTAGAGAGGACGCCCTGCTCCAGGCCCCCGGGGGCGG 540
QY 541 CACAGGACCTGAGGCTCAACTTGTGCGAGGTGCGGTGGAGGCCCAAGGCCCTGCG 600
DB 541 CACAGGACCTGAGGCTCAACTTGTGCGAGGTGCGGTGGAGGCCCAAGGCCCTGCG 600
QY 601 CTAGAGCGCTGCTCTGAGGGCTGCGCGCGCTTCTCAATTGCCAGCTTCAAGGAGCTG 660
DB 601 CTAGAGCGCTGCTCTGAGGGCTGCGCGCGCTTCTCAATTGCCAGCTTCAAGGAGCTG 660
QY 661 GAGCAGCGCTGAGACACCCCGTGACGGGCGAGCCAGCCAGCTTCTCTCATC 720
DB 661 GAGCAGCGCTGAGACACCCCGTGACGGGCGAGCCAGCCAGCTTCTCTCATC 720
QY 721 TCTACTGGGGTGAAGATCGGACAGAGATCCGCAAGATCAGCGACTGCTTCACTGC 780
DB 721 TCTACTGGGGTGAAGATCGGACAGAGATCCGCAAGATCAGCGACTGCTTCACTGC 780
QY 781 CACGCTTCCCGTTCTGACGAGGAGGCGCGCTCGGGCCCTGACGAGCTGCAAA 840
DB 781 CACGCTTCCCGTTCTGACGAGGAGGCGCGCTCGGGCCCTGACGAGCTGCAAA 840
QY 841 CAGCAGACGAGCTGACGAGGTCTCGGGAGACAGAGCGTTCAGGCCAGGTG 900
DB 841 CAGCAGACGAGCTGACGAGGTCTCGGGAGACAGAGCGTTCAGGCCAGGTG 900
QY 901 CTAGGCGGGTGTGAGCTGCTGCGGACAGGCGAGGTGCAAGATGAAGGC 960
DB 901 CTAGGCGGGTGTGAGCTGCTGCGGACAGGCGAGGTGCAAGATGAAGGC 960
QY 961 GTGTACTGGCCCTGAACAGGTGACGCTGAGCACCAAGTGCCTCATTTGCCGAG 1020
DB 961 GTGTACTGGCCCTGAACAGGTGACGCTGAGCACCAAGTGCCTCATTTGCCGAG 1020
QY 1021 GCTGTGTCTGTGCGAGACTGCGCCGCTGCGAGAGCCCTGCGGACAGCTCGATG 1080
DB 1021 GCTGTGTCTGTGCGAGACTGCGCCGCTGCGAGAGCCCTGCGGACAGCTCGATG 1080
QY 1081 GAGGAGGAGTGAAGTCCGCTGCTACCGCATCCCTGCGGGACATGCCCCCACTC 1140
DB 1081 GAGGAGGAGTGAAGTCCGCTGCTACCGCATCCCTGCGGGACATGCCCCCACTC 1140
QY 1141 ATCCGACCAACCGCTTCAAGGCGCTGAGGAGTGCCTGCTACGCGGTGGG 1200
DB 1141 ATCCGACCAACCGCTTCAAGGCGCTGAGGAGTGCCTGCTACGCGGTGGG 1200
QY 1201 CGCTACAGAGGTCAACCCCGCTCCCTACACATCATCACTTCCCTTCTGTTGCT 1260
DB 1201 CGCTACAGAGGTCAACCCCGCTCCCTACACATCATCACTTCCCTTCTGTTGCT 1260
QY 1261 GTGATGTTGGGGATGTTGGGACAGGCTGCTCATGTTCTTTCGCTTGGCCATGCTC 1320
DB 1261 GTGATGTTGGGGATGTTGGGACAGGCTGCTCATGTTCTTTCGCTTGGCCATGCTC 1320
QY 1321 CTTGGCGAGAACGACCGGCTGTGAAGCGCGGAGAACGAGATCTGGCAGACTTTCTTC 1380

DB 1321 CTTGGCGAGAACGACCGGCTGTGAAGCGCGGAGAACGAGATCTGGCAGACTTTCTTC 1380
QY 1381 AGGGCGCGTACTGCTCTGCTTATGAGCCTGTTCTCCATACACCGGCTTCACTAC 1440
DB 1381 AGGGCGCGTACTGCTCTGCTTATGAGCCTGTTCTCCATCTACACCGGCTTCACTAC 1440
QY 1441 AACGAGTGTTCAGTTCGGCCACGAGCATCTTCCCTCGGGCTGAGTGTGGCGCATG 1500
DB 1441 AACGAGTGTTCAGTTCGGCCACGAGCATCTTCCCTCGGGCTGAGTGTGGCGCATG 1500
QY 1501 GCAACACAGTCTGGCTGGAGTGAATCTTCTGGCCAGCACACAGATGCTTACCTGGAT 1560
DB 1501 GCAACACAGTCTGGCTGGAGTGAATCTTCTGGCCAGCACACAGATGCTTACCTGGAT 1560
QY 1561 CCAACAGTCAACCGTGTCTTCTGGAGCCCTAACCCCTTTGGCATCGATCTATTGGAGC 1620
DB 1561 CCAACAGTCAACCGTGTCTTCTGGAGCCCTAACCCCTTTGGCATCGATCTATTGGAGC 1620
QY 1621 CTGGCTGCCAACCACTTGAAGTCTCACTCTCTCAAGATGAAGTGTCCGTCACTCCTG 1680
DB 1621 CTGGCTGCCAACCACTTGAAGTCTCTCACTCTCTCAAGATGAAGTGTCCGTCACTCCTG 1680
QY 1681 GCGCTGCTGACATGCGCTTTGGGCTGCTCTCGAGTCTTCAACAGCTGCACTTTGGC 1740
DB 1681 GCGCTGCTGACATGCGCTTTGGGCTGCTCTCGAGTCTTCAACAGCTGCACTTTGGC 1740
QY 1741 CAGAGCACCGGCTGCTGCTGAGAGCGTGCAGGAGCTCACTTCTGCTGGAGCTTTC 1800
DB 1741 CAGAGCACCGGCTGCTGCTGAGAGCGTGCAGGAGCTCACTTCTGCTGGAGCTTTC 1800
QY 1801 GGTACTCTGCTTCTTAGTCTATCAAGTGGCTGTGTGTGGCTGCGAGGCGCGC 1860
DB 1801 GGTACTCTGCTTCTTAGTCTATCAAGTGGCTGTGTGTGGCTGCGAGGCGCGC 1860
QY 1861 TCGCCACAGATCCTCATCCACTTCAACATGTTCTTCTTCCACAGACCCAGCAAC 1920
DB 1861 TCGCCACAGATCCTCATCCACTTCAACATGTTCTTCTTCCACAGACCCAGCAAC 1920
QY 1921 AGGCTGCTTACCCCGGAGAGGTGTCAGGCGACGCTGCTGTGCTTGGCCTTGGC 1980
DB 1921 AGGCTGCTTACCCCGGAGAGGTGTCAGGCGACGCTGCTGTGCTTGGCCTTGGC 1980
QY 1981 ATGCTGCTTACCCCGGAGAGGTGTCAGGCGACGCTGCTGTGCTTGGCCTTGGC 2040
DB 1981 ATGCTGCTTACCCCGGAGAGGTGTCAGGCGACGCTGCTGTGCTTGGCCTTGGC 2040
QY 2041 CTGCGAGAGGAGGCGCGCTGACCGACAGGAGGAGGAGGCGGCTTGTGGACCTGCT 2100
DB 2041 CTGCGAGAGGAGGCGCGCTGACCGACAGGAGGAGGAGGCGGCTTGTGGACCTGCT 2100
QY 2101 GACGATCTGTGAATGGCTGGAGCTCGATGAGGAGGAGGAGGCGGCTTGTGATGAG 2160
DB 2101 GACGATCTGTGAATGGCTGGAGCTCGATGAGGAGGAGGAGGCGGCTTGTGATGAG 2160
QY 2161 GAGGAGGCGAGCTGCTCCCTCGAGGTCTATGACCGGCGCATCCACACCTCAG 2220
DB 2161 GAGGAGGCGAGCTGCTCCCTCGAGGTCTATGACCGGCGCATCCACACCTCAG 2220
QY 2221 TTTCTGCTGGGCTGCTTCCAAACCGCTCTTACCTGCGCTTGTGGGCTTGGAGCTG 2280
DB 2221 TTTCTGCTGGGCTGCTTCCAAACCGCTCTTACCTGCGCTTGTGGGCTTGGAGCTG 2280
QY 2281 GCGCACCGCAGCTGTGCGAGGTTCTGCGGCGATGCTGATGCGCATAGGCGCTGGCCTG 2340
DB 2281 GCGCACCGCAGCTGTGCGAGGTTCTGCGGCGATGCTGATGCGCATAGGCGCTGGCCTG 2340
QY 2341 GCGCGGAGGTTGGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
DB 2341 GCGCGGAGGTTGGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
QY 2401 ATGAGCGTGGCTTATCTGCTGTGTGATGAGGAGTCTCAGCCTTCTTGCACGCTTGC 2460

```
Db 2401 ATGACCGTGGCTATCCCTGCTGGTATGAGGGGACTCTCAGGCTTCCTGCGACGCCCTCGGG 2460
Qy 2461 CTGCACTGGGTGGAATTCAGAAACAAGTTTCTACTCAGGCAAGGCTCAAGCTGAGTCCC 2520
Db 2461 CTGCACTGGGTGGAATTCAGAAACAAGTTTCTACTCAGGCAAGGCTCAAGCTGAGTCCC 2520
Qy 2521 TTCAACCTTCGCTGCGACAGATGACTAGGCGCCACTGCAAGTTCCTGCGAGACCTCTTCCT 2580
Db 2521 TTCAACCTTCGCTGCGACAGATGACTAGGCGCCACTGCAAGTTCCTGCGAGACCTCTTCCT 2580
Qy 2581 GACCTCTGAGCGAGGAGGAATAAAGACGGTCCGCGCCCTGGCAAAAAA 2640
Db 2581 GACCTCTGAGCGAGGAGGAATAAAGACGGTCCGCGCCCTGGCAAAAAA 2640

RESULT 2
AR577091
LOCUS 2640 bp DNA linear PAT 14-DEC-2004
DEFINITION Sequence 1 from patent US 677537.
ACCESSION AR577091
VERSION AR577091.1 GI:56579543
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2640)
AUTHORS Stashenko, P. and Li, Y.-P.
TITLE Oteoclast proton pump subunit
JOURNAL Patent: US 677537-A 1 17-AUG-2004;
Forsyth Dental Infirmary for Children; Boston, MA;
WOX;

FEATURES
source Location/Qualifiers
1..2640
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 100.0%; Score 2640; DB 2; Length 2640;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGCGTGGCGGACGGGACGCCAGCAGCGAGGCGGCGGCGAGCACACCCGGGGACCATG 60
Db 1 CGGCGTGGCGGACGGGACGCCAGCAGCGAGGCGGCGGCGAGCACACCCGGGGACCATG 60
Qy 61 GGCTCCATGTTCCGGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCACAGCGGCT 120
Db 61 GGCTCCATGTTCCGGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCACAGCGGCT 120
Qy 121 GCTACACCTGCTGAGTGGCTGGGAGCTGGGCTCTGGAGTTTCAGAGACCTCAAC 180
Db 121 GCTACACCTGCTGAGTGGCTGGGAGCTGGGCTCTGGAGTTTCAGAGACCTCAAC 180
Qy 181 GCCTCGGTGAGCGCTTCAGAGACGCTTTGCTGTGATGTTTGGCGCTGTGAGGAGCTG 240
Db 181 GCCTCGGTGAGCGCTTCAGAGACGCTTTGCTGTGATGTTTGGCGCTGTGAGGAGCTG 240
Qy 241 GAGAAGACCTTCACTTCCTGCGAGGAGGAGTGGCGGCGCTGGGCTGGTCTGCCCGCG 300
Db 241 GAGAAGACCTTCACTTCCTGCGAGGAGGAGTGGCGGCGCTGGGCTGGTCTGCCCGCG 300
Qy 301 CCAGAGGAGGCTGCGGACCCCAACCCCGGACCTGTGCGATTCAGAGAGGAGCG 360
Db 301 CCAGAGGAGGCTGCGGACCCCAACCCCGGACCTGTGCGATTCAGAGAGGAGCG 360
Qy 361 GAGCGCTGGCCAGGAGCTGCGGATGTGCGGGCAACCCAGAGCGCTGCGGGCCAG 420
Db 361 GAGCGCTGGCCAGGAGCTGCGGATGTGCGGGCAACCCAGAGCGCTGCGGGCCAG 420
Qy 421 CTGCAACAGCTGAGCTCCACCGCGCGTGTCTAGCGCGAGGCCATGAACCTCAGCTGGCA 480
Db 421 CTGCAACAGCTGAGCTCCACCGCGCGTGTCTAGCGCGAGGCCATGAACCTCAGCTGGCA 480
```

```
Qy 481 GCCGCCCAACACAGATGGGGCTCAGAGAGAGCGCCCTGTCTCAGGCCCCCGGGGGCGG 540
Db 481 GCCGCCCAACACAGATGGGGCTCAGAGAGAGCGCCCTGTCTCAGGCCCCCGGGGGCGG 540
Qy 541 CACCAGGACCTGAGGCTCAACTTTGTGGCAGGTGCGCTGGAGCCCAAGGCCCCCTGCC 600
Db 541 CACCAGGACCTGAGGCTCAACTTTGTGGCAGGTGCGCTGGAGCCCAAGGCCCCCTGCC 600
Qy 601 CTAGAGCGCTGTCTCTGGAGGGCTCGCGGGCTTCTCTCATTTGCCAGCTTCAGGAGCTG 660
Db 601 CTAGAGCGCTGTCTCTGGAGGGCTCGCGGGCTTCTCTCATTTGCCAGCTTCAGGAGCTG 660
Qy 661 GAGCAGCGCTGAGGACACCCCGTGAACCGGCGAGCGACGACGCTGGATGACCTTCTCATC 720
Db 661 GAGCAGCGCTGAGGACACCCCGTGAACCGGCGAGCGACGACGCTGGATGACCTTCTCATC 720
Qy 721 TCCTACTGGGGTGGAGAGATCGGACAAAGATCCGCAAGATTCGGAAGATTCGGAAGTTC 780
Db 721 TCCTACTGGGGTGGAGAGATCGGACAAAGATTCGGAAGATTCGGAAGTTCGGAAGTTC 780
Qy 781 CACGTCTTCCCGTTTCTGACAGGAGGAGGCGCCCTCGGGGCGCTTCGAGCAGCTGCAA 840
Db 781 CACGTCTTCCCGTTTCTGACAGGAGGAGGAGGCGCCCTCGGGGCGCTTCGAGCAGCTGCAA 840
Qy 841 CAGCAGGCGCAGGAGCTGAGGAGGCTCTCGGGGAGACAGAGCGGTTCTTGAGCCAGTG 900
Db 841 CAGCAGGCGCAGGAGCTGAGGAGGCTCTCGGGGAGACAGAGCGGTTCTTGAGCCAGTG 900
Qy 901 CTAGGCGGGTGTCTGAGCTGCTGCGCCAGGCGAGTGCAGGTTCACAGATGAAGGCC 960
Db 901 CTAGGCGGGTGTCTGAGCTGCTGCGCCAGGCGAGTGCAGGTTCACAGATGAAGGCC 960
Qy 961 GTGTACTTGGCCCTCTGAAACAGTGCAGCGTGAGCACCAACGACAGTGCCTCATTCGCGAG 1020
Db 961 GTGTACTTGGCCCTCTGAAACAGTGCAGCGTGAGCACCAACGACAGTGCCTCATTCGCGAG 1020
Qy 1021 GCCTGTGCTCTGTGAGAGCTGCGCGCCCTGAGAGAGGCGCTGCGGAGCAGCTCGATG 1080
Db 1021 GCCTGTGCTCTGTGAGAGCTGCGCGCCCTGAGAGAGGCGCTGCGGAGCAGCTCGATG 1080
Qy 1081 GAGGAGGAGTGTGCTGCGGCTCACCGCATCCCTGCGGGGACATGCCCCCAGACTC 1140
Db 1081 GAGGAGGAGTGTGCTGCGGCTCACCGCATCCCTGCGGGGACATGCCCCCAGACTC 1140
Qy 1141 ATCCGACCAACCGCTTTCAGCGGCTTCAGCGGATCGTGGATCGTACGCGTGGG 1200
Db 1141 ATCCGACCAACCGCTTTCAGCGGCTTCAGCGGATCGTGGATCGTACGCGTGGG 1200
Qy 1201 CGCTACAGAGGTCAACCCCGCTCCCTACCATCATCACTTCCCTTCTGTTTGTCT 1260
Db 1201 CGCTACAGAGGTCAACCCCGCTCCCTACCATCATCACTTCCCTTCTGTTTGTCT 1260
Qy 1261 GTGATGTTGGGATGTGGGCGACGGGCTGCTCATGTTCTCTTGGCCCTGGCCATGGTC 1320
Db 1261 GTGATGTTGGGATGTGGGCGACGGGCTGCTCATGTTCTCTTGGCCCTGGCCATGGTC 1320
Qy 1321 CTTTGGGAGAACCGACCGGCTGTGAAGCGCGCAGAACAGATCTGGCAGACTTCTTC 1380
Db 1321 CTTTGGGAGAACCGACCGGCTGTGAAGCGCGCAGAACAGATCTGGCAGACTTCTTC 1380
Qy 1381 AGGGGCGCTTACCTGCTCTGTTATGGGCTGTTCTCATCTACACCGGCTCATCTAC 1440
Db 1381 AGGGGCGCTTACCTGCTCTGTTATGGGCTGTTCTCATCTACACCGGCTCATCTAC 1440
Qy 1441 AACGAGTGTTCAGTCCGCGCACAGCATCTTCCCTCGGGCTGGAGTGGCGCCCATG 1500
Db 1441 AACGAGTGTTCAGTCCGCGCACAGCATCTTCCCTCGGGCTGGAGTGGCGCCCATG 1500
Qy 1501 GCCAACAGTCTGGCTGGAGTGTGATTCCTGGCGCCAGCACAGATGCTTACCTCGGAT 1560
Db 1501 GCCAACAGTCTGGCTGGAGTGTGATTCCTGGCGCCAGCACAGATGCTTACCTCGGAT 1560
Qy 1561 CCAACGCTACCGGCTGTCTTCTGGGAGCCCTACCCCTTTGGCATCGATCTATTTGGAGC 1620
```

1561 CCACAGCTCACCGGTGTCTTCTGGGACCTTACCCCTTTGGCATCGATCTATTGGAGC 1620
1621 CTGGCTGCCAACACCTTGAGCTCTCAACTCTTCAAGATGAAGATGTCCTGATCTCCG 1680
1621 CTGGCTGCCAACACCTTGAGCTTCTCAACTCTTCAAGATGAAGATGTCCTGATCTCCG 1680
1681 GGCGTCTGTCACATGCGCTTTGGGGTGTCTCGAGTCTTCAACACGCTGCACTTTGGC 1740
1681 GGCGTCTGTCACATGCGCTTTGGGGTGTCTCGAGTCTTCAACACGCTGCACTTTGGC 1740
1741 CAGAGCACCGGTCTGCTGAGAGCGTCCCGAGCTCACCTTCTGCTGGAGCTTTC 1800
1741 CAGAGCACCGGTCTGCTGAGAGCGTCCCGAGCTCACCTTCTGCTGGAGCTTTC 1800
1801 GGTTACCTCGTCTTCTAGTCTATCAAGTGGCTGTGTCTGGCTGCGAGGCGCGC 1860
1801 GGTTACCTCGTCTTCTAGTCTATCAAGTGGCTGTGTCTGGCTGCGAGGCGCGC 1860
1861 TCGCCAGCATCTCATCCACTTCAACATGTTCTCTTCTCCACAGCCCGACAGCAAC 1920
1861 TCGCCAGCATCTCATCCACTTCAACATGTTCTCTTCTCCACAGCCCGACAGCAAC 1920
1921 AGGCTGCTTACCCCGGAGAGGTGTCCAGGCGACGCTGTGTCTGCTGGCTTGGCC 1980
1921 AGGCTGCTTACCCCGGAGAGGTGTCCAGGCGACGCTGTGTCTGCTGGCTTGGCC 1980
1981 ATGTGCCCCATCTGCTGTGTGGCACACCCCTGCACTGCTGCACCGCCACCGCGCGC 2040
1981 ATGTGCCCCATCTGCTGTGTGGCACACCCCTGCACTGCTGCACCGCCACCGCGCGC 2040
2041 CTGCGAGAGGCGCGCTGACCGACAGGAGGAAAAAAGGCGGGTGTGTGACCTGCTCT 2100
2041 CTGCGAGAGGCGCGCTGACCGACAGGAGGAAAAAAGGCGGGTGTGTGACCTGCTCT 2100
2101 GACGATCTGTGAATGGCTGGAGTCCGATGAGGAAAAAGGAGGGGCTGTGATGATGAA 2160
2101 GACGATCTGTGAATGGCTGGAGTCCGATGAGGAAAAAGGAGGGGCTGTGATGATGAA 2160
2161 GAGGAGGCGAGCTGTGCTCCCTCCGAGGTGCTCATGCAACAGGCGCTTGGCGCTGAG 2220
2161 GAGGAGGCGAGCTGTGCTCCCTCCGAGGTGCTCATGCAACAGGCGCTTGGCGCTGAG 2220
2221 TTCTGCTGGCTGTGCTTCCACACCGCTCTTACCTGCGCTGTGGCGCTGAGCGCTG 2280
2221 TTCTGCTGGCTGTGCTTCCACACCGCTCTTACCTGCGCTGTGGCGCTGAGCGCTG 2280
2281 GCCACGCCCGAGCTGTCCGAGGTTCTGTGGGCCATGTTGATGCGCATAGGCTTGGCGCTG 2340
2281 GCCACGCCCGAGCTGTCCGAGGTTCTGTGGGCCATGTTGATGCGCATAGGCTTGGCGCTG 2340
2341 GGCCGGGAGGTGGCGGTGGCGCTGTGTGCTGCTGCTCCCATCTTTGGCGCTTGGCGGTG 2400
2341 GGCCGGGAGGTGGCGGTGGCGCTGTGTGCTGCTGCTCCCATCTTTGGCGCTTGGCGGTG 2400
2401 ATGACCGTGGCTATCTGCTGTGTGATGAGGAGCTCTCAGCGCTTCTGCAAGCGCTTGGCG 2460
2401 ATGACCGTGGCTATCTGCTGTGTGATGAGGAGCTCTCAGCGCTTCTGCAAGCGCTTGGCG 2460
2461 CTGCACTGGGTGGAAATTCAGAAACAGTTCTACTAGGCAAGGCTTCAAGCTGAGTCCC 2520
2461 CTGCACTGGGTGGAAATTCAGAAACAGTTCTACTAGGCAAGGCTTCAAGCTGAGTCCC 2520
2521 TTCACTCTCGCTGCCACAGTACCTAGGCGCCACTGCGAGTCTTCCAGACCTCTTCTCT 2580
2521 TTCACTCTCGCTGCCACAGTACCTAGGCGCCACTGCGAGTCTTCCAGACCTCTTCTCT 2580
2581 GACCTCTGAGGAGGAGGAAATAAGACGGTCCGCGCTTGGCAAAAAAAGGAGGAGGAGG 2640
2581 GACCTCTGAGGAGGAGGAAATAAGACGGTCCGCGCTTGGCAAAAAAAGGAGGAGGAGG 2640

AR270531
LOCUS AR270531 2655 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 1094 from patent US 6500938.
ACCESSION AR270531
VERSION AR270531.1 GI:29701765
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2655)
AUTHORS Au-Young, J. and Seilhamer, J.J.
TITLE Composition for the detection of signaling pathway gene expression
JOURNAL Patent: US 6500938-A 1094 31-DEC-2002;
Incyte Genomics, Inc.; Palo Alto, CA;
WOX;
FEATURES
source Location/Qualifiers
1..2655
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 100.0%; Score 2640; DB 2; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGCGCTGCGCGAGCGGACGAGCGAGCGGCGGCGAGCACACCCGGGGACCATG 60
DB 1 CGCGCTGCGCGAGCGGACGAGCGAGCGGCGGCGAGCACACCCGGGGACCATG 60
QY 61 GGTCCATGTTCCGAGGAGGAGTGGCCCTGTGTCAGCTCTTTTGTGCCACAGCGGT 120
DB 61 GGTCCATGTTCCGAGGAGGAGTGGCCCTGTGTCAGCTCTTTTGTGCCACAGCGGT 120
QY 121 GCCTACACTGCGTGTGAGTGGCTGGCGAGCTGGGCTCTGTTGAGTGTGAGGAGCTG 180
DB 121 GCCTACACTGCGTGTGAGTGGCTGGCGAGCTGGGCTCTGTTGAGTGTGAGGAGCTG 180
QY 181 GCTCGGTGAGCGCTTCCAGAGAGCTTTGTGTTGATGTTGTTGCGCTGTGAGGAGCTG 240
DB 181 GCTCGGTGAGCGCTTCCAGAGAGCTTTGTGTTGATGTTGTTGCGCTGTGAGGAGCTG 240
QY 241 GAGAAAGCTTTACCTTCTCTGAGGAGAGGTGCGCGGCTGGGCTGTGCTCTGCCCCG 300
DB 241 GAGAAAGCTTTACCTTCTCTGAGGAGAGGTGCGCGGCTGGGCTGTGCTCTGCCCCG 300
QY 301 CCAAGGGAGGCTGCGGCGACCCCAACCGGAGCTGCTGCGCATCCAGAGAGAGCG 360
DB 301 CCAAGGGAGGCTGCGGCGACCCCAACCGGAGCTGCTGCGCATCCAGAGAGAGCG 360
QY 361 GAGCGCTGGCCCCAGGAGCTGCGGATGTCGGGGCAACAGAGGCGCTTGGCGGCCAG 420
DB 361 GAGCGCTGGCCCCAGGAGCTGCGGATGTCGGGGCAACAGAGGCGCTTGGCGGCCAG 420
QY 421 CTGCAACAGCTGCACTCCACGCGCCCTGTCTAGCCAGGGCCATGAACCTCAGCTGGCA 480
DB 421 CTGCAACAGCTGCACTCCACGCGCCCTGTCTAGCCAGGGCCATGAACCTCAGCTGGCA 480
QY 481 GCGGCCCCACAGATGGGGCTTCAGAGAGAGCGCCCTGCTCCAGGCCCCCGGGGGCGG 540
DB 481 GCGGCCCCACAGATGGGGCTTCAGAGAGAGCGCCCTGCTCCAGGCCCCCGGGGGCGG 540
QY 541 CACAGGACCTGAGGGTCAAATTTGTGGCAGGTGCGGTGGAGGCCCAAGGCGCCCTGCG 600
DB 541 CACAGGACCTGAGGGTCAAATTTGTGGCAGGTGCGGTGGAGGCCCAAGGCGCCCTGCG 600
QY 601 CTAGAGCGCTGCTCTGAGGCGCTGCGCGGCTTCTCAATTCAGCTTTCAGGAGCTG 660
DB 601 CTAGAGCGCTGCTCTGAGGCGCTGCGCGGCTTCTCAATTCAGCTTTCAGGAGCTG 660
QY 661 GAGCAGCGCTGGAGCACCCCGTGAACGGCGAGCCAGCCAGCTGATGACCTTCTCATC 720
DB 661 GAGCAGCGCTGGAGCACCCCGTGAACGGCGAGCCAGCCAGCTGATGACCTTCTCATC 720

QY 721 TCCTACTGGGGTGGAGCATCGGACAGAGATCCGAGATCATCGGACTGCTTCCACTGC 780
Db 721 TCCTACTGGGGTGGAGCATCGGACAGAGATCCGAGATCATCGGACTGCTTCCACTGC 780
QY 781 CAGCTCTTCCGGTTTCTGACAGAGAGAGAGCCCGCTTCGGGGCCCTGACAGCTGCA 840
Db 781 CAGCTCTTCCGGTTTCTGACAGAGAGAGAGCCCGCTTCGGGGCCCTGACAGCTGCA 840
QY 841 CAGCAGAGCCAGAGCTGACAGAGTCTTCCGGGAGACAGAGCGGTTCTGAGCCAGGTG 900
Db 841 CAGCAGAGCCAGAGCTGACAGAGTCTTCCGGGAGACAGAGCGGTTCTGAGCCAGGTG 900
QY 901 CTAGCCCGGGTGTCTGACGTCTGCTCCGACAGGAGGTGCAAGTCCCAAGATCAAGGCC 960
Db 901 CTAGCCCGGGTGTCTGACGTCTGCTCCGACAGGAGGTGCAAGTCCCAAGATCAAGGCC 960
QY 961 GTGTACTGCTGCTGACAGAGTCCGAGTGTGACGATCCAGCAAGTGTCTTATGCGGAG 1020
Db 961 GTGTACTGCTGCTGACAGAGTCCGAGTGTGACGATCCAGCAAGTGTCTTATGCGGAG 1020
QY 1021 GCCTGTGCTCTGTGCGAGACTGCGCGCTTCCGAGAGGCGCTTCCGAGGACAGCTCGATG 1080
Db 1021 GCCTGTGCTCTGTGCGAGACTGCGCGCTTCCGAGAGGCGCTTCCGAGGACAGCTCGATG 1080
QY 1081 GAGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 1081 GAGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
QY 1141 ATCCGACCAACCGCTTTCAGGCGCAGCTTTCAGGCGCATCGTGTGATCGTACGCGGTGGC 1200
Db 1141 ATCCGACCAACCGCTTTCAGGCGCAGCTTTCAGGCGCATCGTGTGATCGTACGCGGTGGC 1200
QY 1201 CGCTACAGAGAGTCAACCGCGCTTTCAGGCGCATCGTGTGATCGTACGCGGTGGC 1260
Db 1201 CGCTACAGAGAGTCAACCGCGCTTTCAGGCGCATCGTGTGATCGTACGCGGTGGC 1260
QY 1261 GTGATGTTCCGGGATGTGGGCGCAGGCGTGTGATGTTCTTCTTCCGCTTCTTCTTCT 1320
Db 1261 GTGATGTTCCGGGATGTGGGCGCAGGCGTGTGATGTTCTTCTTCCGCTTCTTCTTCT 1320
QY 1321 CTTGCGGAGAACCGACCGCTTGTGAAAGCGCGCAGAGATCTGCGAGACTTCTTCTTCT 1380
Db 1321 CTTGCGGAGAACCGACCGCTTGTGAAAGCGCGCAGAGATCTGCGAGACTTCTTCTTCT 1380
QY 1381 AGGGCCGCTTACCTGCTCTGCTTATGGGCGCTGTTCTTCAATCAACCGGCTTCAATCA 1440
Db 1381 AGGGCCGCTTACCTGCTCTGCTTATGGGCGCTGTTCTTCAATCAACCGGCTTCAATCA 1440
QY 1441 AACGAGTGTCTTCACTGCGGCGCAGCAGATCTTCCCTCGGCTGGAGTGGCGGCGCATG 1500
Db 1441 AACGAGTGTCTTCACTGCGGCGCAGCAGATCTTCCCTCGGCTGGAGTGGCGGCGCATG 1500
QY 1501 GCCAACCACTGTGCTGAGTGTGATGATCTTCTGCGCCAGCAGCAGATGTTACCTCGAT 1560
Db 1501 GCCAACCACTGTGCTGAGTGTGATGATCTTCTGCGCCAGCAGCAGATGTTACCTCGAT 1560
QY 1561 CCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
Db 1561 CCCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620
QY 1621 CTGCTGCGCAACCACTTGTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
Db 1621 CTGCTGCGCAACCACTTGTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
QY 1681 GGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Db 1681 GGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
QY 1741 CAGAGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
Db 1741 CAGAGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
QY 1801 GGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860

Db 1801 GGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
QY 1861 TCGCCAGCATCTTCCATCAATCAATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1920
Db 1861 TCGCCAGCATCTTCCATCAATCAATGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1920
QY 1921 AGGCTGCTTACCCCGGCGAGAGTGTTCAGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
Db 1921 AGGCTGCTTACCCCGGCGAGAGTGTTCAGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
QY 1981 ATGCTGCTTACCCCGGCGAGAGTGTTCAGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
Db 1981 ATGCTGCTTACCCCGGCGAGAGTGTTCAGGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
QY 2041 CTGCGAGGAGCGGCTGACCCGACAGAGGAGAAACAGGCGGCTTCTGAGACTGCTGCTGCTGCT 2100
Db 2041 CTGCGAGGAGCGGCTGACCCGACAGAGGAGAAACAGGCGGCTTCTGAGACTGCTGCTGCTGCT 2100
QY 2101 GACCATCTGTGAATGCTGAGCTCCGATGAGGAGAAAGGAGGCGGCTTCTGAGACTGCTGCTGCT 2160
Db 2101 GACCATCTGTGAATGCTGAGCTCCGATGAGGAGAAAGGAGGCGGCTTCTGAGACTGCTGCTGCT 2160
QY 2161 GAGGAGCGGAGCTGCTCCGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
Db 2161 GAGGAGCGGAGCTGCTCCGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
QY 2221 TTCTG 2280
Db 2221 TTCTG 2280
QY 2281 GCCCAGCGGAGTGTCCGAGGCTTCTGAGGCTTCTGAGGCTTCTGAGGCTTCTGAGGCTTCTG 2340
Db 2281 GCCCAGCGGAGTGTCCGAGGCTTCTGAGGCTTCTGAGGCTTCTGAGGCTTCTGAGGCTTCTG 2340
QY 2341 GCGCGGAGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
Db 2341 GCGCGGAGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
QY 2401 ATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
Db 2401 ATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
QY 2461 CTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
Db 2461 CTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
QY 2521 TTACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 TTACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
QY 2581 GACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
Db 2581 GACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640

RESULT 4

AR380371

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

AR380371

Sequence 916 from patent US 6607879.

AR380371.1 GI:40088005

Unknown.

Unclassified.

1 (bases 1 to 2655)

Cocks, B.G., Stuart, S.G. and Seilhamer, J.J.

Compositions for the detection of blood cell and immunological

response gene expression

Patent: US 6607879-A 916 19-AUG-2003;

Incyte Corporation; Palo Alto, CA

Location/Qualifiers

2655 bp

DNA

linear

PAT 18-DEC-2003

| | | | | |
|----------------------------|------|--|------|--|
| Db | 2041 | CTGCGGAGAGCGCGCTGACCGACAGAGAGAAACAAGGCGCGGTTGCTGAGACTGCCT | 2100 | |
| Qy | 2101 | GACCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAAGGAGGGGGCTGGATGATGAA | 2150 | |
| Db | 2101 | GACGCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAAGGAGGGGGCTGGATGATGAA | 2150 | |
| Qy | 2161 | GAGGAGCGGAGCTGCTCCCTCCGAGGTGCTCATGACACAGGCCATCCACCATCGAG | 2220 | |
| Db | 2161 | GAGGAGCGGAGCTGCTCCCTCCGAGGTGCTCATGACACAGGCCATCCACCATCGAG | 2220 | |
| Qy | 2221 | TTCTGCTGGGCTGCTCTCAACAACCGCCCTCTACCTGCGCCCTGTGGGGCCCTGAGCCTG | 2280 | |
| Db | 2221 | TTCTGCTGGGCTGCTCTCAACAACCGCCCTCTACCTGCGCCCTGTGGGGCCCTGAGCCTG | 2280 | |
| Qy | 2281 | GCCACGCGCCAGCTGTCGAGGTTCTGTGGCCATGTTGATGCGCATAGGCTTGGGCTG | 2340 | |
| Db | 2281 | GCCACGCGCCAGCTGTCGAGGTTCTGTGGCCATGTTGATGCGCATAGGCTTGGGCTG | 2340 | |
| Qy | 2341 | GGCCGGAGGTGGGCGTGGGCGCTGTGGTCTGCTGCCATCTTTGGCGCTTTGGCGTG | 2400 | |
| Db | 2341 | GGCCGGAGGTGGGCGTGGGCGCTGTGGTCTGCTGCCATCTTTGGCGCTTTGGCGTG | 2400 | |
| Qy | 2401 | ATGACCGTGGCTATCTGCTGCTGATGAGGAGCTCTCAGGCTTCTCAGCGCCCTGGG | 2460 | |
| Db | 2401 | ATGACCGTGGCTATCTGCTGCTGATGAGGAGCTCTCAGGCTTCTCAGCGCCCTGGG | 2460 | |
| Qy | 2461 | CTGCACTGGGTGGAATTCAGAAACAAGTTTCTACTCAGGCGGGCTACAGCTGAGTCCC | 2520 | |
| Db | 2461 | CTGCACTGGGTGGAATTCAGAAACAAGTTTCTACTCAGGCGGGCTACAGCTGAGTCCC | 2520 | |
| Qy | 2521 | TTCACTTCCGCTGCACAGATGACTAGGCGCCCACTGCAAGTCTCCTCCAGCTCTCTTCT | 2580 | |
| Db | 2521 | TTCACTTCCGCTGCACAGATGACTAGGCGCCCACTGCAAGTCTCCTCCAGCTCTCTTCT | 2580 | |
| Qy | 2581 | GACCTCTGAGGCGAGGAGGAATAAGACGGTCCGCGCTGGCAAAAAAAGGAAAAA | 2640 | |
| Db | 2581 | GACCTCTGAGGCGAGGAGGAATAAGACGGTCCGCGCTGGCAAAAAAAGGAAAAA | 2640 | |
| RESULT 5 | | | | |
| AX332228 | | | | |
| LOCUS | | | | |
| DEFINITION | | | | |
| SEQUENCE | | | | |
| ACCESSION | | | | |
| VERSION | | | | |
| KEYWORDS | | | | |
| SOURCE | | | | |
| ORGANISM | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | | | | |
| JOURNAL | | | | |
| FEATURES | | | | |
| source | | | | |
| ORIGIN | | | | |
| Query Match | | | | |
| Best Local Similarity | | | | |
| Matches 2640; Conservative | | | | |
| Qy | | | | |
| Db | | | | |

| | | | | |
|----|------|---|------|--|
| Qy | 61 | GGCTCCATGTTTCCGAGAGGAGAGGTGGCTCTGTGTCAGCTCTTTCTGCCCCACAGCGCT | 120 | |
| Db | 61 | GGCTCCATGTTTCCGAGAGGAGAGGTGGCTCTGTGTCAGCTCTTTCTGCCCCACAGCGCT | 120 | |
| Qy | 121 | GCCTTACACCTGCGTGGCTGGGCGAGCTGGGCTCTGAGTTCAGAGACTCAAC | 180 | |
| Db | 121 | GCCTTACACCTGCGTGGCTGGGCGAGCTGGGCTCTGAGTTCAGAGACTCAAC | 180 | |
| Qy | 181 | GCCTCGGTGAGCGCTTCCAGAGACGCTTTGTGTGTTGATGTTGTGGCGCTGTGAGGAGCTG | 240 | |
| Db | 181 | GCCTCGGTGAGCGCTTCCAGAGACGCTTTGTGTGTTGATGTTGTGGCGCTGTGAGGAGCTG | 240 | |
| Qy | 241 | GAGAAGACCTTCACTTCTGTCAGAGAGAGGTGGGCGGCTGGGCTGGTCTGCCCCCG | 300 | |
| Db | 241 | GAGAAGACCTTCACTTCTGTCAGAGAGAGGTGGGCGGCTGGGCTGGTCTGCCCCCG | 300 | |
| Qy | 301 | CCAAAGGGGAGGCTGCGGCGACCCCAACCGCGGACCTGCTGCGCATCCAGAGAGGAGG | 360 | |
| Db | 301 | CCAAAGGGGAGGCTGCGGCGACCCCAACCGCGGACCTGCTGCGCATCCAGAGAGGAGG | 360 | |
| Qy | 361 | GAGCGCTGGCCCAAGGAGCTGCGGGAATGTGCGGGGCAACCAAGAGGCGCTCGCGGCCAG | 420 | |
| Db | 361 | GAGCGCTGGCCCAAGGAGCTGCGGGAATGTGCGGGGCAACCAAGAGGCGCTCGCGGCCAG | 420 | |
| Qy | 421 | CTGCAACAGCTGAGCTCCACGCGCGCTGTACGCCAGGCGCCATGAACCTCAGCTGCGCA | 480 | |
| Db | 421 | CTGCAACAGCTGAGCTCCACGCGCGCTGTACGCCAGGCGCCATGAACCTCAGCTGCGCA | 480 | |
| Qy | 481 | GCGCCCAACAGATGGGCGCTCAGAGAGGAGCGCCCTGCTCCAGGCGCCCGGGGGCGG | 540 | |
| Db | 481 | GCGCCCAACAGATGGGCGCTCAGAGAGGAGCGCCCTGCTCCAGGCGCCCGGGGGCGG | 540 | |
| Qy | 541 | CACCAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGGTGGAGCCCAAGAGCGCCCTGCC | 600 | |
| Db | 541 | CACCAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGGTGGAGCCCAAGAGCGCCCTGCC | 600 | |
| Qy | 601 | CTAGAGCGCTGCTCTGAGAGGCGCTGCGCGGCTTCTCATTTGCGAGGCTTCAAGGAGCTG | 660 | |
| Db | 601 | CTAGAGCGCTGCTCTGAGAGGCGCTGCGCGGCTTCTCATTTGCGAGGCTTCAAGGAGCTG | 660 | |
| Qy | 661 | GAGCAGCGCTGAGGACACCCGCTGACGGGCGAGCGCCAGCGTGGATGACCTTCTCATC | 720 | |
| Db | 661 | GAGCAGCGCTGAGGACACCCGCTGACGGGCGAGCGCCAGCGTGGATGACCTTCTCATC | 720 | |
| Qy | 721 | TCCTACTGGGGTGAAGCAGATCGGACAGAAAGATCCGCAAGATCACGGACTGCTTCACTGC | 780 | |
| Db | 721 | TCCTACTGGGGTGAAGCAGATCGGACAGAAAGATCCGCAAGATCACGGACTGCTTCACTGC | 780 | |
| Qy | 781 | CAGCTCTTCCCGTTTCTGACAGGAGGAGGCGCGCTCGGGGCGCTTCTGAGCAGGTG | 840 | |
| Db | 781 | CAGCTCTTCCCGTTTCTGACAGGAGGAGGCGCGCTCGGGGCGCTTCTGAGCAGGTG | 840 | |
| Qy | 841 | CAGCAGGCGGAGCTGAGGAGGTCTCGGGGAGGAGGCGGTCTTCTGAGCAGGTG | 900 | |
| Db | 841 | CAGCAGGCGGAGCTGAGGAGGTCTCGGGGAGGAGGCGGTCTTCTGAGCAGGTG | 900 | |
| Qy | 901 | CTAGGCGGGTGTGTCAGCTGCTGCGCCAGGCGAGGTGAGGTCCCAAGATGAAGGCC | 960 | |
| Db | 901 | CTAGGCGGGTGTGTCAGCTGCTGCGCCAGGCGAGGTGAGGTCCCAAGATGAAGGCC | 960 | |
| Qy | 961 | GTGTACTGCGCTTGAACCAAGTGTGAGCAACCAAGTGTGAGTGTGCTTCTGAGGAG | 1020 | |
| Db | 961 | GTGTACTGCGCTTGAACCAAGTGTGAGCAACCAAGTGTGAGTGTGCTTCTGAGGAG | 1020 | |
| Qy | 1021 | GCCTGTGCTCTGTGAGGACCTTCCCGCTTCCAGGAGGCGCTTCTGAGGAGCTG | 1080 | |
| Db | 1021 | GCCTGTGCTCTGTGAGGACCTTCCCGCTTCCAGGAGGCGCTTCTGAGGAGCTG | 1080 | |
| Qy | 1081 | GAGGAGGAGTGTGCTGCGGCTCACCGCATCCCTTCCGGGACATGCCCCCACAATC | 1140 | |
| Db | 1081 | GAGGAGGAGTGTGCTGCGGCTCACCGCATCCCTTCCGGGACATGCCCCCACAATC | 1140 | |

Db 2401 ATGACCGTGGCTATCTGCTGGTGATGGAGGACTCTCAGCCTTCTGACGCCCTGCGG 2460
Qy 2461 CTGCACTGGGTGAATTCAGAAACAAGTTCTACTAGGCAAGGGCTTACAAAGCTGAGTCCC 2520
Db 2461 CTGCACTGGGTGAATTCAGAAACAAGTTCTACTAGGCAAGGGCTTACAAAGCTGAGTCCC 2520
Qy 2521 TTCAACCTTCGCTGCCACAGATGACTAGGGCCCACTGCAGGTCTCTGCAGACCTCCCTTCCT 2580
Db 2521 TTCAACCTTCGCTGCCACAGATGACTAGGGCCCACTGCAGGTCTCTGCAGACCTCCCTTCCT 2580
Qy 2581 GACCTCTGAGGAGGAGAGGAATAAAGAGCGTCCGCCCTGGCAAAAAAAGAAAAA 2640
Db 2581 GACCTCTGAGGAGGAGAGGAATAAAGAGCGTCCGCCCTGGCAAAAAAAGAAAAA 2640

RESULT 7
AX410717
LOCUS
DEFINITION
Sequence 3364 from Patent WO0229103.
ACCESSION
AX410717
VERSION
AX410717.1 GI:21443422

KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.

REFERENCE
1
Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J. G.
Gene expression profiles in liver cancer
Patent: WO 0229103-A 3364 11-APR-2002;
JOURNAL
GENE LOGIC INC (US)

FEATURES
source
1..2655
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. U45285"

ORIGIN
Query Match 100.0%; Score 2640; DB 2; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGCGTGGCGGACGGGACGACGACGAGCGGCGGCGGCGAGCACACCGGGGACCATG 60
Db 1 CGGCGTGGCGGACGGGACGACGACGAGCGGCGGCGGCGAGCACACCGGGGACCATG 60
Qy 61 GGTCTCATGTTCGGGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTGCCCCACAGCGGCT 120
Db 61 GGTCTCATGTTCGGGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTGCCCCACAGCGGCT 120
Qy 121 GCCTACACCTGCGTGAAGTGGGCGAGCTGGGCTCTGTGGAGTTTCAGAGACCTCAAC 180
Db 121 GCCTACACCTGCGTGAAGTGGGCGAGCTGGGCTCTGTGGAGTTTCAGAGACCTCAAC 180
Qy 181 GCTCGGTGAGCGCTTCAGAGACGCTTTGTGTGTGATGTTTGGCGCTGTGAGGAGCTG 240
Db 181 GCTCGGTGAGCGCTTCAGAGACGCTTTGTGTGTGATGTTTGGCGCTGTGAGGAGCTG 240
Qy 241 GAGAGACCTTCACCTTCCTGCGAGGAGGAGTGGCGGGCTGGGCTGGTCTGCCCCCG 300
Db 241 GAGAGACCTTCACCTTCCTGCGAGGAGGAGTGGCGGGCTGGGCTGGTCTGCCCCCG 300
Qy 301 CCAAGGGGAGGCTGCGGACCCCCACCGGAGCTCTGTGCGCATCCAGGAGGAGACG 360
Db 301 CCAAGGGGAGGCTGCGGACCCCCACCGGAGCTCTGTGCGCATCCAGGAGGAGACG 360
Qy 361 GAGCGCTGGCCGAGGAGTGGGATGTGGGGCAACAGCAGGCGCTTGGGGCCAG 420
Db 361 GAGCGCTGGCCGAGGAGTGGGATGTGGGGCAACAGCAGGCGCTTGGGGCCAG 420

Qy 421 CTGCAACGAGCTGACGCTCCAGCGCGCTGTACGCCAGGGCCATGAACCTCAGCTGGCA 480
Db 421 CTGCAACGAGCTGACGCTCCAGCGCGCTGTACGCCAGGGCCATGAACCTCAGCTGGCA 480
Qy 481 GCGCGCCACACAGATGGGGCTCTCAGAGGAGCGCCCTGTCTCAGGCCCCCGGGGGCGG 540
Db 481 GCGCGCCACACAGATGGGGCTCTCAGAGGAGCGCCCTGTCTCAGGCCCCCGGGGGCGG 540
Qy 541 CACCAAGACCTGAGGGTCAACTTTGTGTGAGGTGCGGTGGAGCCCAAGGCCCTCGCC 600
Db 541 CACCAAGACCTGAGGGTCAACTTTGTGTGAGGTGCGGTGGAGCCCAAGGCCCTCGCC 600
Qy 601 CTAGAGCGCTCTCTGAGGGGCTGCGCGGCTTCTCATTTGCCAGCTTCAGGAGCTG 660
Db 601 CTAGAGCGCTCTCTGAGGGGCTTGGCGGCTTCTCATTTGCCAGCTTCAGGAGCTG 660
Qy 661 GAGCAGCGCTGGAGCACCCCTGTGACGGGCGAGCCAGCCACGCTGATGACCTTCTCATC 720
Db 661 GAGCAGCGCTGGAGCACCCCTGTGACGGGCGAGCCAGCCACGCTGATGACCTTCTCATC 720
Qy 721 TCCTACTGGGGTGAAGATCGGACAGAGATCCGAAAGATCACGGACTGCTTCCACTGC 780
Db 721 TCCTACTGGGGTGAAGATCGGACAGAGATCCGAAAGATCACGGACTGCTTCCACTGC 780
Qy 781 CACGCTTCCCGTTTCTGCAGCAGGAGGAGGCGCCCTCGGGGCGCTGCAGCAGCTGCAA 840
Db 781 CACGCTTCCCGTTTCTGCAGCAGGAGGAGGCGCCCTCGGGGCGCTGCAGCAGCTGCAA 840
Qy 841 CAGCAGAGCCAGGAGCTGCGAGGAGTCTCTCGGGGAGACAGACGGTCTCTGAGCCAGGTG 900
Db 841 CAGCAGAGCCAGGAGCTGCGAGGAGTCTCTCGGGGAGACAGACGGTCTCTGAGCCAGGTG 900
Qy 901 CTAGGCGGGTGTCTGAGCTGCTGCGGCGAGGAGGTGCGAGTCCAAGATGAAGGCC 960
Db 901 CTAGGCGGGTGTCTGAGCTGCTGCGGCGAGGAGGTGCGAGTCCAAGATGAAGGCC 960
Qy 961 GTGTACTTGGCCCTGAACACAGTGCAGGTGAGCACACGACCAAGTGCCTCATTTGCCAG 1020
Db 961 GTGTACTTGGCCCTGAACACAGTGCAGGTGAGCACACGACCAAGTGCCTCATTTGCCAG 1020
Qy 1021 GCTTGGTGTCTGTGCGAGACCTGCGCCCTGCGAGGAGGCGCTTGGGGACAGCTCGATG 1080
Db 1021 GCTTGGTGTCTGTGCGAGACCTGCGCCCTGCGAGGAGGCGCTTGGGGACAGCTCGATG 1080
Qy 1081 GAGGAGGAGTGAAGTGGCTGCTCACCGCATCCCTGCGGGGACATGCCCCCACCACATC 1140
Db 1081 GAGGAGGAGTGAAGTGGCTGCTCACCGCATCCCTGCGGGGACATGCCCCCACCACATC 1140
Qy 1141 ATCCGACCAACCGCTTCAAGGCGAGCTTCAAGGGCATCGTGGATCGCTACGGCGTGGGC 1200
Db 1141 ATCCGACCAACCGCTTCAAGGCGAGCTTCAAGGGCATCGTGGATCGCTACGGCGTGGGC 1200
Qy 1201 CGCTACAGGAGGTCAACCCCGCTCCCTACACCATCATCCTTCCCTTCTGTTTGTCT 1260
Db 1201 CGCTACAGGAGGTCAACCCCGCTCCCTACACCATCATCCTTCCCTTCTGTTTGTCT 1260
Qy 1261 GTGATGTTGGGGATGTGGGCCACGGGCTGCTCATGTTCTTCCCTCGCCCTGGCCATGCTC 1320
Db 1261 GTGATGTTGGGGATGTGGGCCACGGGCTGCTCATGTTCTTCCCTCGCCCTGGCCATGCTC 1320
Qy 1321 CTTTGGGAGAACCGGAGCGGCTGTGAAAGCGCGGAGAGAGATCTGGCAGACTTTCTTC 1380
Db 1321 CTTTGGGAGAACCGGAGCGGCTGTGAAAGCGCGGAGAGAGATCTGGCAGACTTTCTTC 1380
Qy 1381 AGGGGCGGCTACCTGCTCTGCTTATAGGCGCTGTTCTCACTACACCGGCTTCATCTAC 1440
Db 1381 AGGGGCGGCTACCTGCTCTGCTTATAGGCGCTGTTCTCACTACACCGGCTTCATCTAC 1440
Qy 1441 AACAGTGTCTTCAAGTGGGCGACAGCATCTTCCCTGGGGCTGGAGTGTGGCGCCATG 1500
Db 1441 AACAGTGTCTTCAAGTGGGCGACAGCATCTTCCCTGGGGCTGGAGTGTGGCGCCATG 1500
Qy 1501 GCCAACCAAGTCTGGGTGGAGTGAATGCAATTTCTTGGGCCAGCACAGATGCTTACCTGGAT 1560

[illegible]

Db 2581 GACCTCTGAGCGCAGGAGAGGAATAAAGACGGTCCGCCCTGSCAAACAAAAA 2640

RESULT 8
 LOCUS HSU45285 2655 bp mRNA linear PRI 25-APR-1996
 DEFINITION Human specific 116-kDa vacuolar proton pump subunit (OC-116KDa)
 ACCESSION U45285
 VERSION U45285.1 GI:1245045
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.
 1 (bases 1 to 2655)
 Li,Y.P., Chen,W. and Stashenko,P.
 Molecular cloning and characterization of a putative novel human
 osteoclast-specific 116-kDa vacuolar proton pump subunit
 Biochem. Biophys. Res. Commun. 218 (3), 813-821 (1996)
 PUBMED 8579597
 2 (bases 1 to 2655)
 Li,Y.P., Chen,W. and Stashenko,P.
 Direct Submission
 Submitted (09-JAN-1996) Yi-Ping Li, Cytokine Biology, Forsyth
 Dental Center, 140 Fenway, Boston, MA 02115, USA
 JOURNAL Location/Qualifiers
 FEATURES
 source
 1. 2655
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="osteoclastoma tumor"
 1. 2655
 /gene="OC-116KDa"
 58..2547
 /gene="OC-116KDa"
 /note="novel membrane bound protein with at least six
 transmembrane domains"
 /codon_start=1
 /product="specific 116-kDa vacuolar proton pump subunit"
 /protein_id="AAA97878.1"
 /db_xref="GI:1245046"
 /translations="MGSMFRSEVALVQLFLPAAAYTCVSRIGELGVFEFDRLNASV
 SAQRFRVDVDMRCELEKFTFLQEEVRAGLVLPKGRLPAPPDRLLRIQETE
 RLAQLDRVGNQOALRAQLHQLHLAAVLRQHEPOLAAHTDGAERTPLIQPGG
 PHQDLRVVAGAEAPQHPALERLWRACRFGLTASFRELSQPLEHPVTPGSPATWT
 FLTSYNGEQIKRITDPCFHVFPEFLQBEARLQALQQLQSQEQLQEVGLGETE
 FLSQVLRLQLLPPQOVQHKMAVYALNOCVSTTHKCLIASACSVRLDPLQEE
 NSEEGVSAVAHRI PCRDMPPTLIRNFTASFGQIVDRYGVRYOEVNPAFY
 IITFPLFAVMFGDVGHLMLFALAMVLAENRPAVKAQNEIMQTFPGRYLLILM
 GLFSITGYFYNCEFSRATSI PFSGWSVAMANQSGWSDAFLAQTMLTLDPNVTGV
 LGSPYGPDIPLSLANLHSLFNSFKRMKSVILGVVHMAFVGVLGVFNHVFHGQRRL
 LLDLPETPLTGLFGYLVLYIKWLCVMAAASPILIHFMFLFHSHPGNRLH
 YPQEVQVQATLVALAMVYLIILGTPLHLHRHRRRLRRPADRQENKAGLLDLP
 VPSGWSSEKAGLDDBEEALVPSVLHQAHTIEFLCGVSNVNTASLYRLWALS
 LAHQALSEVLWAMVWIGLGLGREGVAAVLPVIFAAPAVMTVAILLVMEGLSAFLH
 ALRLHWVFQNKYSYGTGKLSDFPTFAATDD"

Query Match 100.0%; Score 2640; DB 5; Length 2655;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGCGTCCGGACGGGACGCCAGCAGCGGACGGCGCGGACAGCACACCCGGGGACCATG 60
 |||||
 Db 1 CGGCGTCCGGACGGGACGGGACGCCAGCAGCGGACGGCGCGGACAGCACACCCGGGGACCATG 60
 |||||

Qy 61 GGCTCCATGTTCCGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCACAGCGGCT 120
 |||||
 Db 61 GGCTCCATGTTCCGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCACAGCGGCT 120
 |||||

Qy 121 GCCTACACCTGCGTGAAGTCGGCTGGGCGAGCTGGGGCTCGTGGAGTTCAGAGACCTCAAC 180
 |||||

Db 121 GCTACACCTGCTGAGTCGGCTGGGAGCTGGGCTCGTGGAGTTCAGAGACCTCAAC 180
Qy 181 GCTCGGTGAGCGCCCTTCAGAGACGCTTTGTTGTTGATGTTTGGCGCTGTGAGGAGCTG 240
Db 181 GCTCGGTGAGCGCCCTTCAGAGACGCTTTGTTGTTGATGTTTGGCGCTGTGAGGAGCTG 240
Qy 241 GAGAAGACCTTACCTTCCTGAGGAGGAGGTCGGCGGGCTGGGCTGCTTCGCCCCCG 300
Db 241 GAGAGACCTTACCTTCCTGAGGAGGAGGTCGGCGGGCTGGGCTGCTTCGCCCCCG 300
Qy 301 CCAAGAGGAGGCTGCGGCACCCCAACCCCGGAGCTGCTGCGCATCCAGGAGGAGAGCG 360
Db 301 CCAAGAGGAGGCTGCGGCACCCCAACCCCGGAGCTGCTGCGCATCCAGGAGGAGAGCG 360
Qy 361 GAGCGCTGGCCAGAGCTGCGGATGTCGCGGGCAACAGAGCGCCCTGCGGGCCAG 420
Db 361 GAGCGCTGGCCAGAGCTGCGGATGTCGCGGGCAACAGAGCGCCCTGCGGGCCAG 420
Qy 421 CTGACACAGCTGAGCTCCAGCGCGCGCTGCTAGCGCAGGGCCATGAACCTCAGCTGGCA 480
Db 421 CTGACACAGCTGAGCTCCAGCGCGCGCTGCTAGCGCAGGGCCATGAACCTCAGCTGGCA 480
Qy 481 GCGGCCACACAGATGGGCTCAGAGAGGACGCCCCCTGCTCCAGGCCCCCGGGGGCGG 540
Db 481 GCGGCCACACAGATGGGCTCAGAGAGGACGCCCCCTGCTCCAGGCCCCCGGGGGCGG 540
Qy 541 CACAGGACCTGAGGCTCAACTTTGTTGGCAGTGGCGTGGAGCCCAAGGCCCCCTGCC 600
Db 541 CACAGGACCTGAGGCTCAACTTTGTTGGCAGTGGCGTGGAGCCCAAGGCCCCCTGCC 600
Qy 601 CTAGAGCGCTGCTCTGAGAGGCTGCGCGCGCTTCTCATTTGCCAGCTTCAGGGAGCTG 660
Db 601 CTAGAGCGCTGCTCTGAGAGGCTGCGCGCGCTTCTCATTTGCCAGCTTCAGGGAGCTG 660
Qy 661 GAGCAGCGCTGAGCACCCTGACGCGGCGAGCCAGCCAGCTGATGACCTTCTCTCATC 720
Db 661 GAGCAGCGCTGAGCACCCTGACGCGGCGAGCCAGCCAGCTGATGACCTTCTCTCATC 720
Qy 721 TCTACTGGGGTGAAGATTCGGAAGAGATCCGCAAGATCCGCAAGATCCGCAAGTCTTCACTGC 780
Db 721 TCTACTGGGGTGAAGATTCGGAAGAGATCCGCAAGATCCGCAAGATCCGCAAGTCTTCACTGC 780
Qy 781 CAGCTCTCCGCTTCTGACAGAGGAGGCGCGCTCGGGCCCTGAGCAGCTGCAAA 840
Db 781 CAGCTCTCCGCTTCTGACAGAGGAGGCGCGCTCGGGCCCTGAGCAGCTGCAAA 840
Qy 841 CAGCAGCAGGAGCTGAGGAGTCTCTCGGGAGACAGAGCGGTTCTGAGCCAGGTTG 900
Db 841 CAGCAGCAGGAGCTGAGGAGTCTCTCGGGAGACAGAGCGGTTCTGAGCCAGGTTG 900
Qy 901 CTAGGCGGGTGTGAGCTGCTGCGCCAGGCGAGGTGAGTCCAGAGATGAAGGCC 960
Db 901 CTAGGCGGGTGTGAGCTGCTGCGCCAGGCGAGGTGAGTCCAGAGATGAAGGCC 960
Qy 961 GTGTACTCGGCTGAAACAGTGCAGCTGAGCACCACACAGTGCCTCATTTGCCGAG 1020
Db 961 GTGTACTCGGCTGAAACAGTGCAGCTGAGCACCACAGTGCCTCATTTGCCGAG 1020
Qy 1021 GCTTGTGTCTGTGCGAGACCTGCGCCCTGCGAGAGGCCCTGCGGAGCAGCTCGATG 1080
Db 1021 GCTTGTGTCTGTGCGAGACCTGCGCCCTGCGAGAGGCCCTGCGGAGCAGCTCGATG 1080
Qy 1081 GAGGAGGAGTGAAGTGCCTGAGCTCAGCGGATCGTGGATCGCTACGCGCTGGGCTGGGC 1140
Db 1081 GAGGAGGAGTGAAGTGCCTGAGCTCAGCGGATCGTGGATCGCTACGCGCTGGGCTGGGC 1140
Qy 1141 ATCCGACCAACCGCTTTCAGCGGACAGCTTTCAGGGGATCGTGGATCGCTACGCGCTGGGC 1200
Db 1141 ATCCGACCAACCGCTTTCAGCGGACAGCTTTCAGGGGATCGTGGATCGCTACGCGCTGGGC 1200
Qy 1201 CGCTACAGGAGGTCAACCCCGCTCCCTTACACATCATCACTTCCCTTCTGTTTGT 1260

Db 1201 CGTACCAGGAGGTCAACCCCGCTCCCTACACCATCATCACCTTCCCTTCTGTTTGTCT 1260
Qy 1261 GTGATGTTGCGGAGTGTGGCCACCGGCTGTCTCATGTTCTCTTTCGCCCTTGCCCATGCTC 1320
Db 1261 GTGATGTTGCGGAGTGTGGCCACCGGCTGTCTCATGTTCTCTTTCGCCCTTGCCCATGCTC 1320
Qy 1321 CTTTGGGAGAACCGACCGGCTGTGAAAGCCGCGAGAACGAGATCTGGCAGATCTTCTTC 1380
Db 1321 CTTTGGGAGAACCGACCGGCTGTGAAAGCCGCGAGAACGAGATCTGGCAGATCTTCTTC 1380
Qy 1381 AGGGGCGCTACTGCTCTGCTTATGGGCTGTGTTCTCATCTACACCGGCTTCATCTAC 1440
Db 1381 AGGGGCGCTACTGCTCTGCTTATGGGCTGTGTTCTCATCTACACCGGCTTCATCTAC 1440
Qy 1441 AACAGTGTCTCAGTCGGGCCACCGAGCATCTTCCCTCGGGCTGGAGTGTGGCGCATG 1500
Db 1441 AACAGTGTCTCAGTCGGGCCACCGAGCATCTTCCCTCGGGCTGGAGTGTGGCGCATG 1500
Qy 1501 GCCAACAGTCTGGCTGGAGTGCATCTTCTGGGCCAGCACAGATGCTTACCTCGAT 1560
Db 1501 GCCAACAGTCTGGCTGGAGTGCATCTTCTGGGCCAGCACAGATGCTTACCTCGAT 1560
Qy 1561 CCCAACGTCACCGGTGTCTTCTCGGAGCCCTACCCCTTTGGCATCGATCTTATTTGGAGC 1620
Db 1561 CCCAACGTCACCGGTGTCTTCTCGGAGCCCTACCCCTTTGGCATCGATCTTATTTGGAGC 1620
Qy 1621 CTGGCTGCCAACCACTTGAAGTCTTCAACTCTCTTCAAGATGAAGATGTCCGTCATCTG 1680
Db 1621 CTGGCTGCCAACCACTTGAAGTCTTCAACTCTCTTCAAGATGAAGATGTCCGTCATCTG 1680
Qy 1681 GCGCTGTGTCACATGCGCTTTGGGGTGTCTCGAGTCTTCAACACAGTGCACCTTGGC 1740
Db 1681 GCGCTGTGTCACATGCGCTTTGGGGTGTCTCGAGTCTTCAACACAGTGCACCTTGGC 1740
Qy 1741 CAGAGCACCGGCTGCTGAGAGCGCTGCGGAGCTCACCTTCTGCTGGAGCTCTTTC 1800
Db 1741 CAGAGCACCGGCTGCTGAGAGCGCTGCGGAGCTCACCTTCTGCTGGAGCTCTTTC 1800
Qy 1801 GGTTACTGCTGTTCCTAGTCACTCAAGTGGCTGTGTCTGGCTGCGCAGGCGCGC 1860
Db 1801 GGTTACTGCTGTTCCTAGTCACTCAAGTGGCTGTGTCTGGCTGCGCAGGCGCGC 1860
Qy 1861 TCGCCAGCATCTCATCCACTTCAACATGTTCTCTTCTCCACAGCCCGCAGCAAC 1920
Db 1861 TCGCCAGCATCTCATCCACTTCAACATGTTCTCTTCTCCACAGCCCGCAGCAAC 1920
Qy 1921 AGGCTGCTCTACCCCGGCGAGGAGTGTCCAGGCCACGCTGGTGGTCTGCGCTTGGCC 1980
Db 1921 AGGCTGCTCTACCCCGGCGAGGAGTGTCCAGGCCACGCTGGTGGTCTGCGCTTGGCC 1980
Qy 1981 ATGGTGGCCATCTGCTGTTGGCACACCCCTGACCTGCTGCAACGCGCACCGCGCGC 2040
Db 1981 ATGGTGGCCATCTGCTGTTGGCACACCCCTGACCTGCTGCAACGCGCACCGCGCGC 2040
Qy 2041 CTTGCGAGGAGCGCGCTGACCGACAGGAGGAAACAAAGCCGGGTTGCTGACCTGCTCT 2100
Db 2041 CTTGCGAGGAGCGCGCTGACCGACAGGAGGAAACAAAGCCGGGTTGCTGACCTGCTCT 2100
Qy 2101 GACGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGAGGAGGCGGCTGGATGATGAA 2160
Db 2101 GACGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGAGGAGGCGGCTGGATGATGAA 2160
Qy 2161 GAGGAGGCGGAGCTGCTCCCTCCGAGGTGTCTATGCAACGAGGCCATCCACACCATCGAG 2220
Db 2161 GAGGAGGCGGAGCTGCTCCCTCCGAGGTGTCTATGCAACGAGGCCATCCACACCATCGAG 2220
Qy 2221 TTTCTGCTGGGCTGCTCTCAACACCGCTCTCTTACCTGCGCTGTGGGCTCTGAGCCTG 2280
Db 2221 TTTCTGCTGGGCTGCTCTCAACACCGCTCTCTTACCTGCGCTGTGGGCTCTGAGCCTG 2280
Qy 2281 GCCACGCGCCAGCTGTCCGAGGTTCTGTGGGCTTCTGTGGCCATGTTGATGCGCATAGGCTTGGCCTG 2340
Db 2281 GCCACGCGCCAGCTGTCCGAGGTTCTGTGGGCTTCTGTGGCCATGTTGATGCGCATAGGCTTGGCCTG 2340

| | | | |
|----|------|--|------|
| Qy | 2341 | GGCCGAGAGTGGCGCTGGCGGTGCTGCTGCTGCCAATCTTTGGCGCCTTTGGCGGTG | 2400 |
| Db | 2341 | GGCCGAGAGTGGCGCTGGCGGTGCTGCTGCTGCCAATCTTTGGCGCCTTTGGCGGTG | 2400 |
| Qy | 2401 | ATGACCGTGCTATCTCGCTGCTCATGGAGGACTCTCAGGCCTTCCTGCAGCCCTCGGG | 2460 |
| Db | 2401 | ATGACCGTGCTATCTCGCTGCTCATGGAGGACTCTCAGGCCTTCCTGCAGCCCTCGGG | 2460 |
| Qy | 2461 | CTGCACCTGGGTGGAAATTCAGAAACAAGTTCTACTCAGGCACGGGCTACAAGCTGAGTCCC | 2520 |
| Db | 2461 | CTGCACCTGGGTGGAAATTCAGAAACAAGTTCTACTCAGGCACGGGCTACAAGCTGAGTCCC | 2520 |
| Qy | 2521 | TTCACCTTCGCTGCCACAGATGACTAGCGGCCCACTGCAGGTCCTGCAGACCTCTCTTCCT | 2580 |
| Db | 2521 | TTCACCTTCGCTGCCACAGATGACTAGCGGCCCACTGCAGGTCCTGCAGACCTCTCTTCCT | 2580 |
| Qy | 2581 | GACCTCTGAGCGAGGAGGAATAAAGACGGTCCGCCCTTGGCAAAAAAAAAAAAAAAA | 2640 |
| Db | 2581 | GACCTCTGAGCGAGGAGGAATAAAGACGGTCCGCCCTTGGCAAAAAAAAAAAAAAAA | 2640 |

| | |
|------------|--|
| RESULT | 9 |
| LOCUS | BC032465 |
| DEFINITION | BC032465 Homo sapiens T-cell, immune regulator 1, ATPase, H+ transporting, lysosomal V0 subunit A3, transcript variant 1, mRNA (CDNA clone MGC:40548 IMAGE:5210733), complete cds. |
| ACCESSION | BC032465 |
| VERSION | BC032465.1 GI:21619617 |
| KEYWORDS | MGC. |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; |

REFERENCE
AUTHORS

1 (bases 1 to 2722)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shennem, C.M., Schuler, G.D., Altshul, S.F., Zeeberg, B., Bueto, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A., Rubin, G.M., Hong, L., Srpapleton, M., Soares, M.B., Bonaldi, M.P., Casavant, T.L., Schetz, T.E., Brownstein, M.J., Usdin, T.B., Tothyluki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Ketterman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.C., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J.J., Schmutz, J., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

CONSTM
TITLE
JOURNAL
PUBMED

| | |
|-----------|---|
| REFERENCE | 12477532 |
| AUTHORS | 2 (bases 1 to 2722) |
| CONSTRM | |
| TITLE | NIH MGC Project Direct Submission |
| JOURNAL | Submitted (06-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA |
| REMARK | NIH-MGC Project URL: http://mgc.nci.nih.gov |
| COMMENT | Contact: MGC help desk Email: cgabs-x@mail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.N.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural |

Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mc@nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Mastello, C., Mastriello, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Skasterip, S., Thomas, P.J., Touchman, J.W.,
Tsongeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: TRAK Plate: 64 Row: j Column: 6.

FEATURES
 source
 series: rmax Plate: 64 Row: j Col:umi: 6.
 Location/Qualifiers
 1. .2722
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="WGC:40548 IMAGE:5210733"
 /tissue_type="lung, Spleen, fetal, pooled"
 /clone_lib="NIH MGC 122"
 /lab_host="DH10B"
 /notes="Vector: pCMV-SPORT6"
 1. .2722
 /gene="TCIRG1"
 /notes="synonyms: TTRC7, ATP6N1C, OC116, Stv1, Vph1, a3, Atp6i, ATP6V0A3, OC-116kDa, OPTB1"
 /db_xref="GeneID:103112"
 /db_xref="HGNC:11647"
 /db_xref="IMGT/GENE-DB:11647"
 /db_xref="MIM:604592"
 102. .2594
 /gene="TCIRG1"
 /codon_start=1
 /product="T-cell, immune regulator 1, isoform a"
 /protein_id="AAH32465.1"
 /db_xref="GI:21619618"
 /db_xref="GeneID:10312"
 /db_xref="HGNC:11647"
 /db_xref="IMGT/GENE-DB:11647"
 /db_xref="MIM:604592"
 /translation="MGSMFRSEVALVQLFLPTAAAYTCVSLGELGVLFERDLNASV
 SFAORFVVRRCCELEKFTFLOEVRRAGLVLPKPPKAPPAPDRLLRIQEBTE
 RLAQELRDVNGQALRALQLOLQHAAVLROGHEPOLAAHTDGCASFTPILOAPGG
 PQDILRVNFAVAGAPHKAPALERLLWRACRGFLIASFRELEQPLEHPVGTGPAWTMT
 FLISWGEQIGQIKRIKTCDFCHVFPFPLQBEARGLAQLOQSQSEQLVEVGTQER
 FLSVLVGLTLPQPGVQNGKAVYIALNCSVSTHKLCILAEACSVRDLPALE
 ALRDSMSEGVSAVHR I CRDMPPTLIRNFTFAS FQGI VDAVGYRGYVNPAPY
 IITFFPLFAVMDGVGHLLMFLFALAMVLAENRPAVKAQNEI WQTFPRGYLLLLM
 GLFPIYGTIYNECSRATS I FPGSVMASMANQSGSDAFLAQHTMLTLDPNVTGVF
 LGPYPGIDPIINLSAHLNLSFLNSFKMKSIVLGVVHMAFGVILGVNHHVFGQRRL
 LLETLPULTFLGLGFLVFLVLYIKWLCVMAARAASAFSILHFMFLFHSHPNRL
 LYPREVQATLIVLALAMWP ILLIGTFLPHLHRRRLRRPADRCQENKAGLLDLP
 DASVNGSDKEAGLDQDEEBAELVPSELMHQAHITIEFCLGQSVNTASYLRMLW
 SLHAQLSVLWAMVNR IGLGAGREGVAAVVLVPIFAAFVMTVAILLVMEGLSAFL
 HALRIHWBFONKFYSGTGYKISPTFEATDP"
 CDS

ORIGIN

| | Query Match | 81.8%; | Score 2159; | DB 5; | Length 2722; |
|----|-----------------------|---|--------------|------------|----------------------|
| | Best Local Similarity | 99.7%; | Pred. No. 0; | | |
| | Matches 2599; | Conservative | 0; | Mismatches | 4; Indels 3; Gaps 1; |
| Qy | 38 | GCGCAGCACACCGGGGACCATGGGCTCCATGTTCCGAGCGAGAGGTGGCCCTGGTCC | 97 | | |
| | | | | | |
| Db | 82 | GCGCAGCACACCGGGGACCATGGGCTCCATGTTCCGAGCGAGAGGTGGCCCTGGTCC | 141 | | |
| | | | | | |
| Qy | 98 | AGCTCTTTTCGCCACAGCGGGTGCCTACACCTGCGTGAGTCGGCTGGCGGAGCTGGGCC | 157 | | |
| | | | | | |
| Db | 142 | AGCTCTTTTCGCCACAGCGGGTGCCTACACCTGCGTGAGTCGGCTGGCGGAGCTGGGCC | 201 | | |
| | | | | | |

158 TCCTGGAGTTACAGACCTCAACGCTCGGTGAGCGCTTTCAGAGACGCTTTGTGGTTG 217
Db
202 TCCTGGAGTTACAGACCTCAACGCTCGGTGAGCGCTTTCAGAGACGCTTTGTGGTTG 261
QY
218 ATGTTTGGCGCTGTGAGGAGCTGGAGAACCTTCACTTCTGTCAGGAGGAGTTCGCGC 277
Db
262 ATGTTTCGGCGCTGTGAGGAGCTGGAGAACCTTCACTTCTGTCAGGAGGAGTTCGCGC 321
QY
278 GGGCTGGGCTGGTTCCTGCCCCCGCCAAAGGGAGGCTGGCGACCCCAACCCCGGAGCC 337
Db
322 GGGCTGGGCTGGTTCCTGCCCCCGCCAAAGGGAGGCTGGCGACCCCAACCCCGGAGCC 381
QY
338 TGTGCGCATCCAGGAGGAGACGGAGCGCTGGCCCAAGAGCTGCGGGATGTGCGGGCA 397
Db
382 TGTGCGCATCCAGGAGGAGACGGAGCGCTGGCCCAAGAGCTGCGGGATGTGCGGGCA 441
QY
398 ACAGAGCGCTTGGCGGCCAGCTGCAACAGCTGCAGTCCACGCGCGCTGCTACGCC 457
Db
442 ACAGAGCGCTTGGCGGCCAGCTGCAACAGCTGCAGTCCACGCGCGCTGCTACGCC 501
QY
458 AGGCGCATCAACTGAGCTGGCAGCGCCGACACAGATGGGCGCTCAGAGAGACGCCCC 517
Db
502 AGGCGCATCAACTGAGCTGGCAGCGCCGACACAGATGGGCGCTCAGAGAGACGCCCC 561
QY
518 TGTCTCAGGCCCCCGGGGGCGCCACACAGGCTCAACTTTGTGGCAGGTGCGG 577
Db
562 TGTCTCAGGCCCCCGGGGGCGCCACACAGGCTCAACTTTGTGGCAGGTGCGG 621
QY
578 TGGAGCCCCACAAAGGCCCTGGCCCTAGAGCGCTGCTTGGAGGGCTGCGCGCGCTTCC 637
Db
622 TGGAGCCCCACAAAGGCCCTGGCCCTAGAGCGCTGCTTGGAGGGCTGCGCGCGCTTCC 681
QY
638 TCATTGGCAGCTTCAAGGAGCTGGAGCAGCGCTGGAGCACCCCGTGACGGCGAGCCAG 697
Db
682 TCATTGGCAGCTTCAAGGAGCTGGAGCAGCGCTGGAGCACCCCGTGAGCGGCGAGCCAG 741
QY
698 CCAGCTGATGATCTTCCTCATCTCTACTGGGGTGGAGCAGATCGACAGAGATCCGCA 757
Db
742 CCAGCTGATGATCTTCCTCATCTCTACTGGGGTGGAGCAGATCGACAGAGATCCGCA 801
QY
758 AGATCAAGAGCTGCTTCCATCTGCGCAGCTTCCCGTTTCTGAGCAGAGAGGCGCGCC 817
Db
802 AGATCAAGAGCTGCTTCCATCTGCGCAGCTTCCCGTTTCTGAGCAGAGAGGCGCGCC 861
QY
818 TCGGGGCGCTGACAGAGCTGCAACAGCAGAGCCAGGAGCTGAGGAGGTCCTCGGGAGA 877
Db
862 TCGGGGCGCTGACAGAGCTGCAACAGCAGAGCCAGGAGCTGAGGAGGTCCTCGGGAGA 921
QY
878 CAGAGCGGTTCTGAGCCAGGTGCTAGGCGGGTGTGTCAGCTGTGTCGCGCAGGCGCAGG 937
Db
922 CAGAGCGGTTCTGAGCCAGGTGCTAGGCGGGTGTGTCAGCTGTGTCGCGCAGGCGCAGG 981
QY
938 TGCAGGTCAACAAGATGAAGCGGTGTACTCGGCCCTGAACAGTGCAGCTGAGAGCA 997
Db
982 TGCAGGTCAACAAGATGAAGCGGTGTACTCGGCCCTGAACAGTGCAGCTGAGAGCA 1041
QY
998 CGCACAGTGCCTCATTTGCGAGGCGTGTGCTGTGCGAGACCTGCGCGCGCTGCAGG 1057
Db
1042 CGCACAGTGCCTCATTTGCGAGGCGTGTGCTGTGCGAGACCTGCGCGCGCTGCAGG 1101
QY
1058 AGGCGCTTGGGACAGCTCGATGGAGGAGGAGTGAAGTCCGCTGCTCAACCGCATCCCT 1117
Db
1102 AGGCGCTTGGGACAGCTCGATGGAGGAGGAGTGAAGTCCGCTGCTCAACCGCATCCCT 1161
QY
1118 GCGGGACATGCCCCCACAACATTCGACCAACACCGCTTCAAGGCGAGCTTCCAGGGCA 1177
Db
1162 GCGGGACATGCCCCCACAACATTCGACCAACACCGCTTCAAGGCGAGCTTCCAGGGCA 1221
QY
1178 TGTGATCGCTACGCGCTGGGCGCTACACAGAGGTCAACCGCTTCCCTACACATCA 1237
Db
1222 TGTGATCGCTACGCGCTGGGCGCTACACAGAGGTCAACCGCTTCCCTACACATCA 1281
QY
1238 TCACCTTCCCCTTCTGTTTGTGTGAATGTCGGGAGTGTGGGCCACGCGGCTGCTCATGT 1297

1282 TCACCTTCCCCTTCTGTTTGTGTGAATGTCGGGAGTGTGGGCCACGCGGCTGCTCATGT 1341
QY
1298 TCCTCTTCCGCTTGCCTCATGTGCTTTCGGAGAACCGACCGGCTGTGAAGCGCGGACAG 1357
Db
1342 TCCTCTTCCGCTTGCCTCATGTGCTTTCGGAGAACCGACCGGCTGTGAAGCGCGGACAG 1401
QY
1358 ACGAGATCTGGCAGACTTCTTTCAGGGGCGCTACTGCTCTGCTTATGGGCGCTGTCT 1417
Db
1402 ACGAGATCTGGCAGACTTCTTTCAGGGGCGCTACTGCTCTGCTTATGGGCGCTGTCT 1461
QY
1418 CCATCTTACACCGGCTTCTATCTACAAAGAGTGTCTCAGTCGCGCCACACAGCATCTTCCCCT 1477
Db
1462 CCATCTTACACCGGCTTCTATCTACAAAGAGTGTCTCAGTCGCGCCACACAGCATCTTCCCCT 1521
QY
1478 CGGGCTGAGTGTGCGCCCATGCGCAACAGTCTGGCTGGAGTGAATGATTCCTGGGCC 1537
Db
1522 CGGGCTGAGTGTGCGCCCATGCGCAACAGTCTGGCTGGAGTGAATGATTCCTGGGCC 1581
QY
1538 AGCACAGATGCTTACCTTGGATCCAAAGGTCAACGCTGTCTTCTTGGGACCTTACCCCT 1597
Db
1582 AGCACAGATGCTTACCTTGGATCCAAAGGTCAACGCTGTCTTCTTGGGACCTTACCCCT 1641
QY
1598 TTGGCATCGATCTTATTTGGAGCTTGGCTGCCAACACTTGGAGCTTCTCAACTCTCTTCA 1657
Db
1642 TTGGCATCGATCTTATTTGGAGCTTGGCTGCCAACACTTGGAGCTTCTCAACTCTCTTCA 1701
QY
1658 AGATGAAGATGTCGTCTATCTTGGGCTGTGTCACATGGCTTTGGGCTGTCTCGGAG 1717
Db
1702 AGATGAAGATGTCGTCTATCTTGGGCTGTGTCACATGGCTTTGGGCTGTCTCGGAG 1761
QY
1718 TCCTTCAACACAGTGCATTTGGCCAGAGGCAACGCTGTCTGTCGAGACGCTGCGGAGC 1777
Db
1762 TCCTTCAACACAGTGCATTTGGCCAGAGGCAACGCTGTCTGTCGAGACGCTGCGGAGC 1821
QY
1778 TCACCTTCTCTGCTGGGACTCTTTCGGTTACCTGCTGTCTTCTAGTCTATCTCAAGTGGCTGT 1837
Db
1822 TCACCTTCTCTGCTGGGACTCTTTCGGTTACCTGCTGTCTTCTAGTCTATCTCAAGTGGCTGT 1881
QY
1838 GTGCTGGGCTGTCAGAGGCGCGCTCG---CGCAGATCTCTATCTCACTTCACTCAACATGT 1894
Db
1882 GTGCTGGGCTGTCAGAGGCGCGCTCGGCCCCCAGCATCTCTATCTCACTTCACTCAACATGT 1941
QY
1895 TCCTCTTCTCCACAGCCCGCAGCAACAGCTGTCTTACCCCGCAGAGGAGTGTCCAGG 1954
Db
1942 TCCTCTTCTCCACAGCCCGCAGCAACAGCTGTCTTACCCCGCAGAGGAGTGTCCAGG 2001
QY
1955 CCACGCTGGTGTCTTGGCTTGGCTGTGTCCTTCTGCTGTGGCACAACCCCTGTC 2014
Db
2002 CCACGCTGGTGTCTTGGCTTGGCTGTGTCCTTCTGCTGTGGCACAACCCCTGTC 2061
QY
2015 ACCTGCTGCAACGCAACCGCGCGCTGCGGAGGAGGCGCGCTGACGACAGAGAGAAA 2074
Db
2062 ACCTGCTGCAACGCAACCGCGCGCTGCGGAGGAGGCGCGCTGACGACAGAGAGAAA 2121
QY
2075 ACAAGCGCGGTGTGCTGACCTGCTGACGATCTGTGAATGGCTTGGAGCTCCGATGAGG 2134
Db
2122 ACAAGCGCGGTGTGCTGACCTGCTGACGATCTGTGAATGGCTTGGAGCTCCGATGAGG 2181
QY
2135 AAAAGCGAGGCGCTTGGATGATGAAGAGGAGGCGCGCTGCTGCTTCCCTCCGAGGTGCTCA 2194
Db
2182 AAAAGCGAGGCGCTTGGATGATGAAGAGGAGGCGCGCTGCTGCTTCCCTCCGAGGTGCTCA 2241
QY
2195 TGCACAGGCGCATCCACACCATCGAGTTCGTGCTGGGCTGTGCTTCCAAACCGCGCTCT 2254
Db
2242 TGCACAGGCGCATCCACACCATCGAGTTCGTGCTGGGCTGTGCTTCCAAACCGCGCTCT 2301
QY
2255 ACCTGCGCTGTGGGCGCTGAGCTTGGCCACGCGCAGCTGTCCGAGGTTCGTGGGCGCA 2314
Db
2302 ACCTGCGCTGTGGGCGCTGAGCTTGGCCACGCGCAGCTGTCCGAGGTTCGTGGGCGCA 2361
QY
2315 TGGTGAATCGCATAGGCTTGGGCGCTTGGGCGGAGGTGGGCGTGTGCGGCTGTGCTGTGG 2374

Db 2362 TGGTATGCGCATAGGCTGGGCGCTGGGCGGAGGTGGGCGTGGCTGTGTGCTGG 2421

Qy 2375 TCCCATCTTTGCGGCTTTGCGCGTATGACCGTGGCTATCTCTGTGTGTATGAGGAC 2434

Db 2422 TCCCATCTTTGCGGCGCTTTGCGCGTATGACCGTGGCTATCTCTGTGTATGAGGAC 2481

Qy 2435 TCTCAGGCTTCTCAGCGCTGCGGCTGCACTGGGTGGAATCCAGAACAAAGTCTACT 2494

Db 2482 TCTCAGGCTTCTCAGCGCTGCGGCTGCACTGGGTGGAATCCAGAACAAAGTCTACT 2541

Qy 2495 CAGCAGCGGCTACAAGCTGAGTCCCTTACCTTCGCTGCGCACAGATGACTAGGCGCCAC 2554

Db 2542 CAGCAGCGGCTACAAGCTGAGTCCCTTACCTTCGCTGCGCACAGATGACTAGGCGCCAC 2601

Qy 2555 TGCAGTCTCTGCGCAGACTCTTCTGACCTCTGAGGAGAGAGAGGAGGAGCGGTCC 2614

Db 2602 TGCAGTCTCTGCGCAGACTCTTCTGACCTCTGAGGAGAGAGGAGGAGGAGCGGTCC 2661

Qy 2615 GCCTGGCAAAAAAAAAAAAAAAAAAAAA 2640

Db 2662 GCCTGGCAAAAAAAAAAAAAAAAAAAAA 2687

RESULT 10

BC018133

LOCUS

DEFINITION

Homo sapiens T-cell, immune regulator 1, ATPase, H+ transporting, lysosomal V0 subunit A3, transcript variant 1, mRNA (CDNA clone MGC:9307 IMAGE:3905446), complete cds.

ACCESSION

BC018133

VERSION

BC018133.1 GI:17390297

KEYWORDS

MGC.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 2691)

1 Srausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullah, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., Skalska, U., Marra, M.A.

Mammalian Gene Collection Program Team

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2691)

NIH MGC Project

Direct Submission

Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-x@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 22 Row: i Column: 15

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19924144.

FEATURES

Location/Qualifiers

1..2691

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:9307 IMAGE:3905446"

/tissue_type="Pancreas, epithelioid carcinoma"

/clone_lib="NIH_MGC_70"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1..2691

/gene="TCIRG1"

/notes="synonyms: TIRC7, ATP6N1C, OC116, Stv1, Vph1, a3, Atp6i, ATP6V0A3, OC-116kDa, OPTB1"

/db_xref="GeneID:10312"

/db_xref="HGNC:11647"

/db_xref="IMGT/GENE-DB:11647"

/db_xref="MIM:604592"

109..2601

/gene="TCIRG1"

/codon_start=1

/product="T-cell, immune regulator 1, isoform a"

/protein_id="AAH18133.1"

/db_xref="GI:17390298"

/db_xref="GeneID:10312"

/db_xref="HGNC:11647"

/db_xref="IMGT/GENE-DB:11647"

/db_xref="MIM:604592"

/translations="MSPRSEVALVQLPTAAAYTCVSLGELGLVPRDLNASV SAQRFVVDVRECELEKTFLOEVRAGLVLPKGRLPAPPPRDLRIQBEETE RAQELRDVRNQALRAQLHQLHRAVLRQGHPEQLAAHTDASERTPLQAPGG PHQDLRVNFGVAGVAPKALRACRGLIAFRLEQLPHEPVTGEPATWMT FLISYGEQIGQRIKIDCFCHVFPFLQOEERLQALQOQSELQEVLETTER FLSQVLGRVLQLLPQGVQVHKMAYLALNQCSTVTHKCLIAEAWCSVDLPALOE ALRDSMEEGVSVAHRIPCRDMPPTLIRTNRTASPGQIVDAVGCRYOEVPAPVT IITFPLPAMFGDVGHLLMFLPALAMVLAENRPAPVAKAAQNEIWTQFFRGLLLM GLFSITYGFIYNECFRATSIFFSGWSVAANQSGWSDAFLAQTHTLTDPNVTGVF LGPYFGIDPDIWSLAANHLFSFKMMSVILGVVHMAFGVLGVFNHVFHGGHRL LLETLPETLFLGLFGLVFLVYKWLVCWMAAASAPSILIHFINFLFSHSPNRL LYPRQVVOATLVVLAMVPIILLGTPLLHRRHRRRRPADROENKAGLLDLP DASVNGSSDEEKAGGLDDEEELVPSVLMHQAHTIEFCGCVSNTASYLFLWL SLAHAQLSEVLWAMVRIGLGLRGVGAANVLVPIPAAPAVMTVAILLWEGUSAFEL HAURLHWVEFQNKFYSGTGYKUSPFTFAATDD"

ORIGIN

Query Match 81.7%; Score 2156; DB 5; Length 2691;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2596; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

Qy 38 GCGCAGCACACCCGGGACCATGGGCTCCATGTTCCGAGCGAGGAGGTGGCCCTGTCC 97

Db 89 GCGCAGCACACCCGGGACCATGGGCTCCATGTTCCGAGCGAGGAGGTGGCCCTGTCC 148

Qy 98 AGCTCTTCTGCGCCACAGCGGCTCCCTACACCTCGCTGAGTCCGCTGGCGAGTGGGCC 157

Db 149 AGCTCTTCTGCGCCACAGCGGCTCCCTACACCTCGCTGAGTCCGCTGGCGAGTGGGCC 208

Qy 158 TCGTGGAGTTTCAGAGACCTCAACGCTCGGTGAGCGCTTCCAGAGACGCTTCTGTGGTTG 217

Db 209 TCGTGGAGTTTCAGAGACCTCAACGCTCGGTGAGCGCTTCCAGAGACGCTTCTGTGGTTG 268


```
Db 1449 CGCCATGCGCAACAGTCTGGCTGGAGTATGATTCCTTGGCCGACACAGCATGCTTA 1508
QY 1553 CCTGTATCCCAACGTCACGGTGTCTTCTGGGACCCCTACCCCTTTGGGATCGATCCTTA 1612
Db 1509 CCTGTATCCCAACGTCACGGTGTCTTCTGGGACCCCTACCCCTTTGGGATCGATCCTTA 1568
QY 1613 TTTGGAGCTGGCTGCCAACACTTGAGCTTCTCAACTCTCTTCAAGATGAAGATGTCG 1672
Db 1569 TTTGGAGCTGGCTGCCAACACTTGAGCTTCTCAACTCTCTTCAAGATGAAGATGTCG 1628
QY 1673 TCATCTTGGGCTGCTGTCATATGGCTTTTGGGTGTCTTGGAGTCTTCAACACAGTGC 1732
Db 1629 TCATCTTGGGCTGCTGTCATATGGCTTTTGGGTGTCTTGGAGTCTTCAACACAGTGC 1688
QY 1733 ACTTTGGCAGAGGACCGGCTGCTGTCAGAGCGCTGCGGAGCTCACCTTCTGCTGG 1792
Db 1689 ACTTTGGCAGAGGACCGGCTGCTGTCAGAGCGCTGCGGAGCTCACCTTCTGCTGG 1748
QY 1793 GACTCTTGGTACCTCGTGTCTTCTAGTCTATCTACAGTGGCTGTGTGTGGCTGCCA 1852
Db 1749 GACTCTTGGTACCTCGTGTCTTCTAGTCTATCTACAGTGGCTGTGTGTGGCTGCCA 1808
QY 1853 GGGCCGCTCG --- CCAGCATCTCATCCACTTTCATCAACATGTTCTTCTTCCACCA 1909
Db 1809 GGGCCGCTCGGGCCCGCCAGCATCTCATCCACTTTCATCAACATGTTCTTCTTCCACCA 1868
QY 1910 GCGCCAGCAACAGGCTGCTTACCCCGGAGAGAGTGTCTCAGGCGACGCTGTGGTGC 1969
Db 1869 GCGCCAGCAACAGGCTGCTTACCCCGGAGAGAGTGTCTCAGGCGACGCTGTGGTGC 1928
QY 1970 TGGCTTGGCATGTGTGCCATCTCTGCTTGGCACACCCCTGCACCTGCTGCACCGGC 2029
Db 1929 TGGCTTGGCATGTGTGCCATCTCTGCTTGGCACACCCCTGCACCTGCTGCACCGGC 1988
QY 2030 ACGCCGCTGCTGCGAGAGGCGCGCTGACCGACAGAGGAGAAACAAGCGCGGTTC 2089
Db 1989 ACGCCGCTGCTGCGAGAGGCGCGCTGACCGACAGAGGAGAAACAAGCGCGGTTC 2048
QY 2090 TGGACTGTGCTGACGATCTGTGAATGGCTGGAGCTCCGATGAGGAGAAAGCGAGGGCC 2149
Db 2049 TGGACTGTGCTGACGATCTGTGAATGGCTGGAGCTCCGATGAGGAGAAAGCGAGGGCC 2108
QY 2150 TGGATGATGAGAGGAGGCGGAGCTGCTGCCCTCCGAGTGTCTATGACACGAGGCATCC 2209
Db 2109 TGGATGATGAGAGGAGGCGGAGCTGCTGCCCTCCGAGTGTCTATGACACGAGGCATCC 2168
QY 2210 ACACCATCGAGTTCTGCTTGGCTGCTTCCAAACACCGCTTCTACCTGCGCTGTGG 2269
Db 2169 ACACCATCGAGTTCTGCTTGGCTGCTTCCAAACACCGCTTCTACCTGCGCTGTGG 2228
QY 2270 CCTGAGCTTGGCCACGCGGAGCTGCTGAGGTTCTTGGGCGATGATGCGCATAG 2329
Db 2229 CCTGAGCTTGGCCACGCGGAGCTGCTGAGGTTCTTGGGCGATGATGCGCATAG 2288
QY 2330 GCTGGGCTTGGGCGGAGGTTGGGCTGCTGGGCTGCTGGTGTGGTCTTGGCTTGGCG 2389
Db 2289 GCTGGGCTTGGGCGGAGGTTGGGCTGCTGGGCTGCTGGTGTGGTCTTGGCTTGGCG 2348
QY 2390 CTTTGGCTGATGACCGTGGCTATCTCTGCTGATGAGGAGTCTTCAAGCTTCTTCTGC 2449
Db 2349 CTTTGGCTGATGACCGTGGCTATCTCTGCTGATGAGGAGTCTTCAAGCTTCTTCTGC 2408
QY 2450 AGCCCTGGGCTGACCTGGTGGATTCAGAACAGTCTTACTAGGACGAGGCTACA 2509
Db 2409 AGCCCTGGGCTGACCTGGTGGATTCAGAACAGTCTTACTAGGACGAGGCTACA 2468
QY 2510 AGCTGAGTCCCTTCACTTCTGCTGCGACAGTACGAGTGGGCGGCTGCTGCTGCTG 2569
Db 2469 AGCTGAGTCCCTTCACTTCTGCTGCGACAGTACGAGTGGGCGGCTGCTGCTGCTG 2528
QY 2570 ACCTCTTCTGACCTTCTGAGGAGAGAGGAAATAAGACGGTTCGCGCTTGGCA 2623
```

```
Db 2529 ACTCTCTCTCTGACCTCTGAGGAGGAGGAATAAGACGGTTCGCGCTTGGCA 2582

RESULT 12
A98501
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..2488
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
537..2381
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAB70562.1"
/db_xref="GI:6781602"
/translation="MTFLISYMGEOIGKIRKITDCFHCHVFPFQQEERLQALQOL
QOQSOELQEVLETGRFLOVLRVQLPPGOVVOVHKMKAVYLALNOCVSVTHKCL
TAEACSVRDLPALQEARLSDSMERGVSVAHRIPCRDMPTLIRNRTAFSGQIVD
AYGVRYQYQVNPAPYITITFPFLPAVMGDDVGHLLMFLPALAMVLAENRPAVKAQN
EIWQTFPFRGYLLLLMGLFSIYGTINECFSRATSIFFSGWSVMAANQSGWSDAFL
AQHTMLTLDPNVTGFLGPIYFGIDPIWSLAANHLISFLNFKHMSVILGVVHMAFV
VLGVFNHVFQGRHRLLELTPELTFLGLFGYLVFLVYKWLVCVMAARAASAPSLI
HFINMFLSPNSRLLYPROBVQATLVLAAMVPIILLGLTPLHLHRRHRRRLRR
PADROENKAGLLDLPDASVNGSSDEKAGGLDDEEAEIAPSEVLHQAHTIIEFC
LCQVNTASYLRMLWALSIAHQLSEVLNWMVRIGLIGREVGVAIVLVPFPAFV
MTVAILLVMEGLSFLHALRLHWEVFQNKFYSGTGYKLSFPFTFAATDD"

Query Match 63.9%; Score 1687; DB 2; Length 2488;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2077; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 561 CTTTGTGGCAGTGGCGTGGAGCCCAAGGCCCTGCCCTAGAGCGCTGCTCTGGAG 620
Db 392 CTTTGTGGCAGTGGCGTGGAGCCCAAGGCCCTGCCCTAGAGCGCTGCTCTGGAG 451
QY 621 GGCCTGCGCGGCTTCTCATTTGCCAGCTTCAGGGAGCTGGAGCAGCGCTGGAGCACCC 680
Db 452 GGCCTGCGCGGCTTCTCATTTGCCAGCTTCAGGGAGCTGGAGCAGCGCTGGAGCACCC 511
QY 681 CQTGACGGCGAGCAGCAGCCACGTCGATGACCTTCTCATCTCTACTGGGGTGAGCAGAT 740
Db 512 CQTGACGGCGAGCAGCAGCCACGTCGATGACCTTCTCATCTCTACTGGGGTGAGCAGAT 571
QY 741 CGGACAGAGATCCGCAAGATCACGGACTGTCTTCCACTGCCACGCTTCCCGTTCTGCA 800
Db 572 CGGACAGAGATCCGCAAGATCACGGACTGTCTTCCACTGCCACGCTTCCCGTTCTGCA 631
QY 801 CGAGGAGGAGCGCGCTCGGGGCTCGAGCAGCTGCAACAGCAGCAGCAGCAGGCTGCA 860
Db 632 CGAGGAGGAGCGCGCTCGGGGCTCGAGCAGCTGCAACAGCAGCAGCAGCAGGCTGCA 691
QY 861 GGAGGTCTCTCGGGGAGACAGAGCGTTCCTGAGCCAGGTGCTAGGCCGGGTGCTGCAGCT 920
Db 692 GGAGGTCTCTCGGGGAGACAGAGCGTTCCTGAGCCAGGTGCTAGGCCGGGTGCTGCAGCT 751
QY 921 GCTGCGCGCAGGCGAGGTGCGAGGTCCAAGAATGAAGCGCGTGTACTCGGCGCTTGAACCA 980
Db 752 GCTGCGCGCAGGCGAGGTGCGAGGTCCAAGAATGAAGCGCGTGTACTCGGCGCTTGAACCA 811
```

```
QY 981 GTGAGCGTGGACGACGACCAAGTGCCTCATTTGCCGAGGCCCTGGTGTCTCTGTGGAGA 1040
DB 812 GTGAGCGTGGACGACGACCAAGTGCCTCATTTGCCGAGGCCCTGGTGTCTCTGTGGAGA 871
QY 1041 CCTGCCCGCCCTGACGAGGCGCCCTGCCGAGCAGCTCGATGAGGAGGAGTGAGTGCCT 1100
DB 872 CCTGCCCGCCCTGACGAGGCGCCCTGCCGAGCAGCTCGATGAGGAGGAGTGAGTGCCT 931
QY 1101 GGCTACCGCATCCCTCCGCGGACATGCCCCCACAATCATCCGACCAACCGCTTAC 1160
DB 932 GGCTACCGCATCCCTCCGCGGACATGCCCCCACAATCATCCGACCAACCGCTTAC 991
QY 1161 GGCAGCTTCCAGGCGATCGTGAATCGCTACGCGGTGGCGGCTTACAGAGGTCAACCC 1220
DB 992 GGCAGCTTCCAGGCGATCGTGAATCGCTACGCGGTGGCGGCTTACAGAGGTCAACCC 1051
QY 1221 CGCTCCCTACACCATCATCATCTTCCCTTCCCTGTTGCTGTGTGATGTTGCGGAGTGGG 1280
DB 1052 CGCTCCCTACACCATCATCATCTTCCCTTCCCTGTTGCTGTGTGATGTTGCGGAGTGGG 1111
QY 1281 CCACGGGCTGCTCATGTTCTTCTTTCGCTTCCCTGGCCATGGTCTTTCGCGAGAACCGACCGC 1340
DB 1112 CCACGGGCTGCTCATGTTCTTCTTTCGCTTCCCTGGCCATGGTCTTTCGCGAGAACCGACCGC 1171
QY 1341 TGTGAAGCGCGCAGACGAGATCTGGCAGACTTCTTTCAGGGCCGCTTACCTGCTCCT 1400
DB 1172 TGTGAAGCGCGCAGACGAGATCTGGCAGACTTCTTTCAGGGCCGCTTACCTGCTCCT 1231
QY 1401 GCTTATGGGCTGTTTCTTCTTCTTACACCGCTTCTTCTTACAAAGAGTGTTCAGTCCGC 1460
DB 1232 GCTTATGGGCTGTTTCTTCTTCTTACACCGCTTCTTCTTACAAAGAGTGTTCAGTCCGC 1291
QY 1461 CACAGCATCTTCCCTCGGGCTGGAGTGGCGCCATGGCCCAACAGTCTGGCTGGAG 1520
DB 1292 CACAGCATCTTCCCTCGGGCTGGAGTGGCGCCATGGCCCAACAGTCTGGCTGGAG 1351
QY 1521 TGATGCATTTCTGGCCGACACAGATGCTTACCTGGATCCCAAGTCAACGGTGTCTT 1580
DB 1352 TGATGCATTTCTGGCCGACACAGATGCTTACCTGGATCCCAAGTCAACGGTGTCTT 1411
QY 1581 CCTGGGACCTTACCCCTTTGGCATCGATCTATTGGAGCTGGCTGCCAACCACTTGAG 1640
DB 1412 CCTGGGACCTTACCCCTTTGGCATCGATCTATTGGAGCTGGCTGCCAACCACTTGAG 1471
QY 1641 CTTCTCTCAACTCTTCAAGATGAAGATGTCGGTCACTCTGGGGTCTGTGCACATGGCCTT 1700
DB 1472 CTTCTCTCAACTCTTCAAGATGAAGATGTCGGTCACTCTGGGGTCTGTGCACATGGCCTT 1531
QY 1701 TGGGGTGGTCTCGGAGTCTTCAACGATGCACTTGGCCAGAGGACCGGCTGTGCT 1760
DB 1532 TGGGGTGGTCTCGGAGTCTTCAACGATGCACTTGGCCAGAGGACCGGCTGTGCT 1591
QY 1761 GGAGAGCTGCGCGAGCTCACTTCTGCTGGGACTCTTCCGTTTACCTCGTGTCTTAGT 1820
DB 1592 GGAGAGCTGCGCGAGCTCACTTCTGCTGGGACTCTTCCGTTTACCTCGTGTCTTAGT 1651
QY 1821 CATCTCAAGTGGCTGTGTGTCTGGGTGTCAGGGCGCCCTCG ---CCACAGCATCTCTCAT 1877
DB 1652 CATCTCAAGTGGCTGTGTGTCTGGGTGTCAGGGCGCCCTCG ---CCACAGCATCTCTCAT 1711
QY 1878 CCATCTCATCAATGTTCTTCTTCCGACAGCCCGCAGAACAGGCTGTCTACCCCG 1937
DB 1712 CCATCTCATCAATGTTCTTCTTCCGACAGCCCGCAGAACAGGCTGTCTACCCCG 1771
QY 1938 GCAGGAGTGGTCCAGCCACGCTGTGTGTCTGGCTTGGCCATGTCGCTTCCCTGCT 1997
DB 1772 GCAGGAGTGGTCCAGCCACGCTGTGTGTCTGGCTTGGCCATGTCGCTTCCCTGCT 1831
QY 1998 GCTTGGCACACCCCTGCACTGTGTGACCGCCACCGCGCGCTGCGGAGGAGGCGCCG 2057
DB 1832 GCTTGGCACACCCCTGCACTGTGTGACCGCCACCGCGCGCTGCGGAGGAGGCGCCG 1891
```

```
QY 2058 TGACCGACAGAGGAAAAACAAGCGCGGTTGCTGGACCTGCTGACCGCATCTGTGAATGG 2117
DB 1892 TGACCGACAGAGGAAAAACAAGCGCGGTTGCTGGACCTGCTGACCGCATCTGTGAATGG 1951
QY 2118 CTGAGCTCTCGATGAGGAAAAAGGAGGCGGCTGATGATGAAGAGGAGCCGAGCTCGT 2177
DB 1952 CTGAGCTCTCGATGAGGAAAAAGGAGGCGGCTGATGATGAAGAGGAGCCGAGCTCGT 2011
QY 2178 CCCCTCCGAGTGTCTATGACACAGGCCATCCACACCATCGAGTTCTGCTGGGCTGCGT 2237
DB 2012 CCCCTCCGAGTGTCTATGACACAGGCCATCCACACCATCGAGTTCTGCTGGGCTGCGT 2071
QY 2238 CTCACACACCGCTCTTACCTGCGCTGTGGGCGCTGAGCTGCGCCACCGCCAGCTGTC 2297
DB 2072 CTCACACACCGCTCTTACCTGCGCTGTGGGCGCTGAGCTGCGCCACCGCCAGCTGTC 2131
QY 2298 CGAGTTCTGTGGGCGATGTGATGCGCATAGGCTGGGCTGGGCTGGGAGGTGGCGT 2357
DB 2132 CGAGTTCTGTGGGCGATGTGATGCGCATAGGCTGGGCTGGGCTGGGAGGTGGCGT 2191
QY 2358 GCGGCTGTGTGCTGCTGCCCATCTTTTGGCGCTTTGCCGTGATGACCGTGGCTATCCT 2417
DB 2192 GCGGCTGTGTGCTGCTGCCCATCTTTTGGCGCTTTGCCGTGATGACCGTGGCTATCCT 2251
QY 2418 GCTGCTGATGAGGAGCTCTCAGGCTTCTGCGACGCGCTGGGCTGACCTGGGTTGGAAT 2477
DB 2252 GCTGCTGATGAGGAGCTCTCAGGCTTCTGCGACGCGCTGGGCTGACCTGGGTTGGAAT 2311
QY 2478 CCAGAACAGTTCTTACTCAGGCGAGGCTACAAAGCTGAGTCCCTTACCTTTCGCTGCCAC 2537
DB 2312 CCAGAACAGTTCTTACTCAGGCGAGGCTACAAAGCTGAGTCCCTTACCTTTCGCTGCCAC 2371
QY 2538 AGATGACTAGGCGGCTTCTGAGGCTTCTGCGACGCTTCTTCTGACCTCTGAGGAGGAG 2597
DB 2372 AGATGACTAGGCGGCTTCTGAGGCTTCTGCGACGCTTCTTCTGACCTCTGAGGAGGAG 2431
QY 2598 AGGAATAAGACCGTCCGCGCTTGCACAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
DB 2432 AGGAATAAGACCGTCCGCGCTTGCACAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAG 2474

RESULT 13
AX920850 LOCUS 2488 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1 from Patent EP1350848.
ACCESSION AX920850
VERSION AX920850.1 GI:40214620
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 unclassified sequences.
AUTHORS Utku,N., Gullans,S.R. and Milford,E.L.
TITLE T-cell membrane protein (TIRC7), gene encoding it and antibodies
JOURNAL against it and uses thereof
PATENT Patent: EP 1350848-A 1 08-OCT-2003;
UTKU, Nalan (DE); BRIGHAM AND WOMEN'S HOSPITAL (US)
FEATURES
source Location/Qualifiers
1..2488
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
537..2381
/notes="unnamed protein product"
/codon_start=1
/protein_id="CA02141.1"
/db_xref="GI:40214621"
/translation="MTFLISYMGEGIQKIRKITDCHFHVFPFLOQEEARLALQQL
QQSQBLQVLGETERLSQVLGRVQLLPQGVQVHKMAVYLAHQCSVSTHKCL
IAEAMCSVRDLPALQALRDLSSMEEGSAVAHRIPCRDMPPTLIRNRFVSGFQIVD
AVGGRYQEVNPAPYTIITFPFLFAVMFGDVGHLLMFLPALAMLAENRPAVKAQON
EIWQTFPRGYILLLLGLPSIYTGPIYNCFSPATSIFFPSGWSVAAMANGSGHDAFL
AQHTMLTLPNVTVGLFPGFPYFGIDPINSIAANHLNFLNSFRKMSVILGVVHMAFGV
```



```
QY 1221 CGCTCCCTACACCATCATCACCCTCCCTCCCTGTTGCTGTGATGTTGGGGATGTGGG 1280
Db 1052 CGCTCCCTACACCATCATCACCCTCCCTCCCTGTTGCTGTGATGTTGGGGATGTGGG 1111
QY 1281 CCACGGGCTGCTCATGTTCTCTTTCCGCTCGCCCATGCTCTTCCGAGAAACCGACCGGC 1340
Db 1112 CCACGGGCTGCTCATGTTCTCTTTCCGCTCGCCCATGCTCTTCCGAGAAACCGACCGGC 1171
QY 1341 TGTGAAGCCGCGCAGAACGAGATCTGGCAGACTTTCTTACGGGGCCGCTACCTGCTCCT 1400
Db 1172 TGTGAAGCCGCGCAGAACGAGATCTGGCAGACTTTCTTACGGGGCCGCTACCTGCTCCT 1231
QY 1401 GCTTATGGGCTGTTCTCTCATCTACACCGGCTTCTATCTACAAACGAGTCTTTCAGTCCGCG 1460
Db 1232 GCTTATGGGCTGTTCTCTCATCTACACCGGCTTCTATCTACAAACGAGTCTTTCAGTCCGCG 1291
QY 1461 CACCAGCATCTTCCCTCTGGGCTGGAGTGTGGCGCCCATGTCACCAACAGTCTGGCTGGAG 1520
Db 1292 CACCAGCATCTTCCCTCTGGGCTGGAGTGTGGCGCCCATGTCACCAACAGTCTGGCTGGAG 1351
QY 1521 TGATGCATCTTGGCCGACGACAGATGCTTACCTGGATCCCAACGTCACCGGTGCTT 1580
Db 1352 TGATGCATCTTGGCCGACGACAGATGCTTACCTGGATCCCAACGTCACCGGTGCTT 1411
QY 1581 CCTGGGACCTTACCCTTTGGCATCGATCTTATTTGGAGCTGTGCTGCCAACCACTTGAG 1640
Db 1412 CCTGGGACCTTACCCTTTGGCATCGATCTTATTTGGAGCTGTGCTGCCAACCACTTGAG 1471
QY 1641 CTTCTCTCAACTCTTCAAGATGAAGATGTCCGTCATCTTGGGCGTCTGTGCACATGGCCTT 1700
Db 1472 CTTCTCTCAACTCTTCAAGATGAAGATGTCCGTCATCTTGGGCGTCTGTGCACATGGCCTT 1531
QY 1701 TGGGCTGGTCTCGGAGTCTTCAACAGGTGTCATTTGGCCAGAGGACACCGGCTGCTGCT 1760
Db 1532 TGGGCTGGTCTCGGAGTCTTCAACAGGTGTCATTTGGCCAGAGGACACCGGCTGCTGCT 1591
QY 1761 GGAGACGCTGCGGAGTCTCACTTCTGCTGGGACTCTTGGGTTACCTCGTGTTCCTAGT 1820
Db 1592 GGAGACGCTGCGGAGTCTCACTTCTGCTGGGACTCTTGGGTTACCTCGTGTTCCTAGT 1651
QY 1821 CATCTACAAGTGGCTGTGTGTCTGGGCTGCGAGGCGCGCTCG---CCACGATCCTCAT 1877
Db 1652 CATCTACAAGTGGCTGTGTGTCTGGGCTGCGAGGCGCGCTCGCGCCCGACGATCCTCAT 1711
QY 1878 CCACCTCATCAATGTTCTCTTCTCCACAGCCCGCAGCAACAGGCTGCTCTACCCCGG 1937
Db 1712 CCACCTCATCAATGTTCTCTTCTCCACAGCCCGCAGCAACAGGCTGCTCTACCCCGG 1771
QY 1938 GCAGGAGTGTCCAGGCGCAGCTGTGTGCTTGGGCTTGGCCATGTTGCCATCCTGCT 1997
Db 1772 GCAGGAGTGTCCAGGCGCAGCTGTGTGCTTGGGCTTGGCCATGTTGCCATCCTGCT 1831
QY 1998 GCTTGGCACACCCCTGCACCTGTGTGCACCGCCGCGCGCTGCGGAGGAGGCGCGC 2057
Db 1832 GCTTGGCACACCCCTGCACCTGTGTGCACCGCCGCGCGCTGCGGAGGAGGCGCGC 1891
QY 2058 TGACCGACAGAGAGAAACAAAGCCCGGTTGCTGGACCTGCTGCAACGATCTGTGAATGG 2117
Db 1892 TGACCGACAGAGAGAAACAAAGCCCGGTTGCTGGACCTGCTGCAACGATCTGTGAATGG 1951
QY 2118 CTGGAGCTCCGATGAGGAAGGAGGAGGCGCTGGATGATCAAGAGGAGGCGCGAGCTCGT 2177
Db 1952 CTGGAGCTCCGATGAGGAAGGAGGAGGCGCTGGATGATCAAGAGGAGGCGCGAGCTCGT 2011
QY 2178 CCCCTCCGAGGTGCTCATATGACACAGGCGCATCCACCATCGAGTTCTGCTGGGCTGCTG 2237
Db 2012 CCCCTCCGAGGTGCTCATATGACACAGGCGCATCCACCATCGAGTTCTGCTGGGCTGCTG 2071
QY 2238 CTCGACACCGCTCTCTACCTGCGCTGTGGGCGCTGAGGCTGCGCCACGCGCCAGCTGTC 2297
Db 2072 CTCGACACCGCTCTCTACCTGCGCTGTGGGCGCTGAGGCTGCGCCACGCGCCAGCTGTC 2131
QY 2298 CGAGGTTCTGTGGGCCATGATGATGCGCATAGGCTTGGGCTGGGCGGGAGGTGGGCGT 2357
```

```
Db 2132 CGAGGTTCTGTGGGCCATGTTGATGCGCATAGGCTTGGGCTTGGCCCGGAGGTGGCGT 2191
QY 2358 GGCGGCTGTGTGCTGTGCTCCCATCTTTTGGCGCCTTTTGGCCTGATGACCGTGGCTATCCT 2417
Db 2192 GGCGGCTGTGTGCTGTGCTCCCATCTTTTGGCGCCTTTTGGCCTGATGACCGTGGCTATCCT 2251
QY 2418 GCTGCTGATGGAGGAGCTCTCAGGCTTCTGTCACGCTTCTGGGCTGCACTGGGTGGAAT 2477
Db 2252 GCTGCTGATGGAGGAGCTCTCAGGCTTCTGTCACGCTTCTGGCCTGCGCTGCACTGGGTGGAAT 2311
QY 2478 CCAGAACAAAGTTCTTACTCAGSCAGGCTACAAGCTGAGTCCCTTACCTTTCGCTGCCAC 2537
Db 2312 CCAGAACAAAGTTCTTACTCAGSCAGGCTACAAGCTGAGTCCCTTACCTTTCGCTGCCAC 2371
QY 2538 AGATGACTTAGGGGCCACTTGCAGGTCCTGCCAGACTCTCTTCTTGCACCTCTGAGGCGAG 2597
Db 2372 AGATGACTTAGGGGCCACTTGCAGGTCCTGCCAGACTCTCTTCTTGCACCTCTGAGGCGAG 2431
QY 2598 AGGAATAAAGACGCTCCGCTTGGCAAAAAA 2640
Db 2432 AGGAATAAAGACGCTCCGCTTGGCAAAAAA 2474
```

Search completed: June 30, 2006, 14:14:21
Job time : 14902 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2006, 05:24:06 ; Search time 1622 Seconds
(without alignments)
11348.173 Million cell updates/sec

Title: US-10-783-519-1

Perfect score: 2640

Sequence: 1 cggcgctgcggacggcgag.....gcaaaaaaaaaaaaaa 2640

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 5244920 seqs, 3486124231 residues

Word size : 1

Total number of hits satisfying chosen parameters: 10489196

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*
- 14: Geneseqn2005s:*
- 15: Geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 2640 | 100.0 | 2640 | 2 | AAV04529 Human OC- |
| 2 | 2640 | 100.0 | 2640 | 6 | ABN86735 Human OC- |
| 3 | 2640 | 100.0 | 2640 | 13 | ADR44118 Human 116 |
| 4 | 2640 | 100.0 | 2640 | 14 | ADY92605 Human 116 |
| 5 | 2640 | 100.0 | 2655 | 6 | ABL64400 Stomach c |
| 6 | 2640 | 100.0 | 2655 | 6 | ABL68345 Kidney ca |
| 7 | 2640 | 100.0 | 2655 | 6 | ABK84337 Human cdn |
| 8 | 2640 | 100.0 | 2655 | 6 | ABN96866 Gene #336 |
| 9 | 2640 | 100.0 | 2655 | 10 | ACA56496 Human sig |
| 10 | 2640 | 100.0 | 2655 | 11 | ADI31590 Human cdn |
| 11 | 2640 | 100.0 | 2655 | 12 | ADI56292 Human pol |
| 12 | 2640 | 100.0 | 2655 | 13 | ADS83657 Human lym |
| 13 | 2142 | 81.1 | 2700 | 12 | ADI28819 Human mod |
| 14 | 2142 | 81.1 | 2700 | 12 | ADP10401 Reference |
| 15 | 2142 | 81.1 | 2700 | 13 | ACN39267 Tumour-as |
| 16 | 2090 | 79.2 | 2676 | 13 | ADQ86432 Human tum |
| 17 | 2090 | 79.2 | 2676 | 13 | ADQ83344 Human tum |
| 18 | 2046 | 77.5 | 2493 | 14 | ADV43978 Human psy |

| | | | | | |
|----|------|------|-------|----|--------------------|
| 19 | 2039 | 77.2 | 2676 | 12 | ADQ85289 Human tum |
| 20 | 1742 | 66.0 | 2561 | 12 | ADH22582 CDNA enco |
| 21 | 1687 | 63.9 | 2488 | 2 | AAx24912 T-cell me |
| 22 | 1670 | 63.3 | 2480 | 13 | ACN39268 Tumour-as |
| 23 | 1640 | 62.1 | 2762 | 13 | ACN43472 Human dia |
| 24 | 1636 | 62.0 | 2488 | 2 | AAx24913 T-cell me |
| 25 | 1449 | 54.9 | 1845 | 14 | ADV43979 Human psy |
| 26 | 1304 | 49.4 | 3104 | 6 | ABQ54601 Human ova |
| 27 | 612 | 23.2 | 612 | 10 | ADD27224 Human adi |
| 28 | 442 | 16.7 | 493 | 6 | ABL68797 Kidney ca |
| 29 | 442 | 16.7 | 493 | 6 | ABL82832 Human ova |
| 30 | 442 | 16.7 | 493 | 6 | ABN94793 Gene #129 |
| 31 | 373 | 14.1 | 373 | 10 | ADD27651 Human adi |
| 32 | 350 | 13.3 | 425 | 9 | ACH49178 Human teu |
| 33 | 302 | 11.4 | 434 | 6 | ABL81465 Human ova |
| 34 | 288 | 10.9 | 414 | 9 | ACH29799 Human tes |
| 35 | 238 | 9.0 | 238 | 4 | AAI29251 Colon tum |
| 36 | 238 | 9.0 | 238 | 8 | ABZ33437 Human col |
| 37 | 221 | 8.4 | 15941 | 10 | ADC64674 Human TCI |
| 38 | 179 | 6.8 | 189 | 12 | ACH92650 Human gen |
| 39 | 179 | 6.8 | 532 | 12 | ACH78950 Human gen |
| 40 | 179 | 6.8 | 3269 | 4 | AAK69047 Human imm |
| 41 | 176 | 6.7 | 570 | 6 | ABL80219 Human ova |
| 42 | 166 | 6.3 | 419 | 6 | ABL81517 Human ova |
| 43 | 138 | 5.2 | 561 | 6 | ABL82854 Human ova |
| 44 | 124 | 4.7 | 568 | 13 | ADU14483 Solid tum |
| 45 | 110 | 4.2 | 151 | 2 | AAQ72536 Osteoclas |

ALIGNMENTS

RESULT 1

AAV04529 standard; cDNA; 2640 BP.

XX AAV04529;

AC AAV04529;

XX 02-JUL-1998 (first entry)

XX Human OC-116 kDa cDNA sequence.

XX Osteoclast; human; OC-116 kDa; screening; cell surface marker; probe; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 58..2523

XX FT /*tag= a

XX FT /product= "OC-116 kDa"

XX FT /note= "no stop codon given"

XX PN WO9803651-AL.

XX PD 29-JAN-1998.

XX PF 10-JUL-1997; 97WO-US012569.

XX PR 19-JUL-1996; 96US-00684932.

XX PA (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.

XX PI Stashenko P, Li Y, Wucherpfennig AL;

XX DR WPI; 1998-120776/11.

XX DR P-PSDB; AAW41943.

XX Osteoclast specific or related DNA sequence - useful as probe to screen genomic DNA or cDNA library, or as osteoclast cell surface marker.

XX Example 8; Page 58-61; 75pp; English.

XX The present sequence represents the human OC-116 kDa cDNA sequence which

|||||
541 CACACGACCTGAGGCTCAACTTTGTGGAGGTGCGTGGAGCCCAAGGCCCTGGC 600
|||||
601 CTAGAGGGCTGCTCTGAGGGGCTGCCGGCTTCTCATTTGCCAGCTTCAGGAGCTG 660
|||||
601 CTAGAGGGCTGCTCTGAGGGGCTGCCGGCTTCTCATTTGCCAGCTTCAGGAGCTG 660
|||||
661 GAGCAGCGCTGAGACACCCCGTGAAGGGCGAGCCAGCCACAGTGGATGACCTTCTTCATC 720
|||||
661 GAGCAGCGCTGAGACACCCCGTGAAGGGCGAGCCAGCCACAGTGGATGACCTTCTTCATC 720
|||||
721 TCTTACTGGGGTGAAGAGATCGGACAGAAAGATCCGCAAGATCAAGACTGCTTCCAATGC 780
|||||
721 TCTTACTGGGGTGAAGAGATCGGACAGAAAGATCCGCAAGATCAAGACTGCTTCCAATGC 780
|||||
781 CAGCTCTTCCCGTTCTGACAGAGGAGGCGCGCTCGGGCCCTGACAGAGCTGCA 840
|||||
781 CAGCTCTTCCCGTTCTGACAGAGGAGGCGCGCTCGGGCCCTGACAGAGCTGCA 840
|||||
841 CAGCAGGACAGAGCTGACAGAGGTCTCTCGGGAGACAGAGCGGTCTCTGAGCCAGGTG 900
|||||
841 CAGCAGGACAGAGCTGACAGAGGTCTCTCGGGAGACAGAGCGGTCTCTGAGCCAGGTG 900
|||||
901 CTAGGCGGGTGTGAGCTGCTGCGCCAGGCGAGGTGCAAGATCAAGATGAAAGGC 960
|||||
901 CTAGGCGGGTGTGAGCTGCTGCGCCAGGCGAGGTGCAAGATCAAGATGAAAGGC 960
|||||
961 GTGTACTCGGCTGAACAGGTGACGGTGAACACAGCACAAGTGCCTCATTTGCCGAG 1020
|||||
961 GTGTACTCGGCTGAACAGGTGACGGTGAACACAGCACAAGTGCCTCATTTGCCGAG 1020
|||||
1021 GCCTGTGTCTGTGCGAGACCTGCGCCCTGCGAGAGCCCTGCGGACAGCTCGATG 1080
|||||
1021 GCCTGTGTCTGTGCGAGACCTGCGCCCTGCGAGAGCCCTGCGGAGAGCTCGATG 1080
|||||
1081 GAGGAGGAGTGAAGTCCGTGCTACCGCATCCCTGCGGGAATGCCCCCAACTC 1140
|||||
1081 GAGGAGGAGTGAAGTCCGTGCTACCGCATCCCTGCGGGAATGCCCCCAACTC 1140
|||||
1141 ATCCGACCAACCGCTTACGGCCAGCTTCCAGGGCATCGTGGATCGTACGGCTGGGC 1200
|||||
1141 ATCCGACCAACCGCTTACGGCCAGCTTCCAGGGCATCGTGGATCGTACGGCTGGGC 1200
|||||
1201 CGCTACAGAGGTCAACCCCGCTCCCTACACCATCATCACCTTCCCTTTCTGTGTGCT 1260
|||||
1201 CGCTACAGAGGTCAACCCCGCTCCCTACACCATCATCACCTTCCCTTTCTGTGTGCT 1260
|||||
1261 GTGATGTTTGGGATGTGGGCAAGGGCTGCTCATGTTCTTTCGCCCTTGGCCATGGTC 1320
|||||
1261 GTGATGTTTGGGATGTGGGCAAGGGCTGCTCATGTTCTTTCGCCCTTGGCCATGGTC 1320
|||||
1321 CTTTGGGAGAACCGACCGCTGTGAAGCGCGCAGAACGAGATCTGGCAGACTTTCTTC 1380
|||||
1321 CTTTGGGAGAACCGACCGCTGTGAAGCGCGCAGAACGAGATCTGGCAGACTTTCTTC 1380
|||||
1381 AGGGGCGGTACTGCTCTCTGCTTATGGGCTGTTCTTCCATCTACACCGGCTTCACTAC 1440
|||||
1381 AGGGGCGGTACTGCTCTCTGCTTATGGGCTGTTCTTCCATCTACACCGGCTTCACTAC 1440
|||||
1441 AACAGTGTTCAGTTCGGCCACAGCATCTTCCCTCGGGCTGGAGTGGCGGCATG 1500
|||||
1441 AACAGTGTTCAGTTCGGCCACAGCATCTTCCCTCGGGCTGGAGTGGCGGCATG 1500
|||||
1501 GCCAACAGTCTGGCTGGAGTGAATGCAATTCCTGGCCAGCACAGATGCTTACCTCGGAT 1560
|||||
1501 GCCAACAGTCTGGCTGGAGTGAATGCAATTCCTGGCCAGCACAGATGCTTACCTCGGAT 1560
|||||
1561 CCCAAGCTCACCGGTGCTTCTTGGGACCTTACCCCTTTGGCATCGATCTATTTGGAGC 1620
|||||
1561 CCCAAGCTCACCGGTGCTTCTTGGGACCTTACCCCTTTGGCATCGATCTATTTGGAGC 1620
|||||
1621 CTGGCTGCCAACCACTTGAAGTCTCTCAACTCTTCAAGATGAAGATGTCGCTCATCTTG 1680
|||||

Db 1621 CTGGCTGCCAACCACTTGAAGTCTCTCAACTCTTCAAGATGAAGATGTCGCTCATCTTG 1680
Qy 1681 GCGCTCGTGCAATGCGCTTTGGGGTGGTCTCTCGAGTCTTCAACCACTGCACTTTGGC 1740
Db 1681 GCGCTCGTGCAATGCGCTTTGGGGTGGTCTCTCGAGTCTTCAACCACTGCACTTTGGC 1740
Qy 1741 CAGAGCACCGGCTGCTGCTGGAGACGCTGCGCGAGCTCACCTTCTCTGCTGGGACTTCTTC 1800
Db 1741 CAGAGCACCGGCTGCTGCTGGAGACGCTGCGCGAGCTCACCTTCTCTGCTGGGACTTCTTC 1800
Qy 1801 GGTACTCTGTTTCTTCTAGTCAATCAAGTGGCTGTGTCTGGGCTGCCAGGGCGGC 1860
Db 1801 GGTACTCTGTTTCTTCTAGTCAATCAAGTGGCTGTGTCTGGGCTGCCAGGGCGGC 1860
Qy 1861 TCGCCAGCATCCTCATCCACTTCAACATGTTTCTTCTTCCACAGCCCCAGCAAC 1920
Db 1861 TCGCCAGCATCCTCATCCACTTCAACATGTTTCTTCTTCCACAGCCCCAGCAAC 1920
Qy 1921 AGGCTGCTTACCCCGGAGAGGTGTTCCAGGGCCACGCTGCTGCTTGGCTTTGGCC 1980
Db 1921 AGGCTGCTTACCCCGGAGAGGTGTTCCAGGGCCACGCTGCTGCTTGGCTTTGGCC 1980
Qy 1981 ATGTTGCCCATCTGCTGCTTGGCACAACCTGCTGCTGCAACCGCCACCGCCCGCC 2040
Db 1981 ATGTTGCCCATCTGCTGCTTGGCACAACCTGCTGCTGCAACCGCCACCGCCCGCC 2040
Qy 2041 CTGCGGAGGAGGCGCGCTGACCGACAGAGGAGAAACAAGCCGGGTTGCTGGACCTGCGCT 2100
Db 2041 CTGCGGAGGAGGCGCGCTGACCGACAGAGGAGAAACAAGCCGGGTTGCTGGACCTGCGCT 2100
Qy 2101 GACGATCTGTGAATGGCTGGAGCTCCGATGAGAGAAAGAGAGGGGCTTGGATGAGAA 2160
Db 2101 GACGATCTGTGAATGGCTGGAGCTCCGATGAGAGAAAGAGAGGGGCTTGGATGAGAA 2160
Qy 2161 GAGGAGGCGAGCTGCTCCCTCCAGAGTGTCTATGCAACAGGCCATTCACACCATTCGAG 2220
Db 2161 GAGGAGGCGAGCTGCTCCCTCCAGAGTGTCTATGCAACAGGCCATTCACACCATTCGAG 2220
Qy 2221 TTTCTGCTGGGCTGCTGCTCCAAACCGCTCTCTACCTGCGCTGTGGGCTTGGAGCTG 2280
Db 2221 TTTCTGCTGGGCTGCTGCTCCAAACCGCTCTCTACCTGCGCTGTGGGCTTGGAGCTG 2280
Qy 2281 GCCACGCGCAGCTGCTCGAGGTTCTGTGGGCGCATGGTGATGGCATAGGCTTGGGCTG 2340
Db 2281 GCCACGCGCAGCTGCTCGAGGTTCTGTGGGCGCATGGTGATGGCATAGGCTTGGGCTG 2340
Qy 2341 GCGCGGAGTGGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
Db 2341 GCGCGGAGTGGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
Qy 2401 ATGACCGTGGCTATCTGCTGCTGATGGAGGAGTCTCAGCCCTTCTGCAAGGCTTGGCG 2460
Db 2401 ATGACCGTGGCTATCTGCTGCTGATGGAGGAGTCTCAGCCCTTCTGCAAGGCTTGGCG 2460
Qy 2461 CTGCACTGGGTGGAATTCAGAACAAAGTTCTACTCAGGCAAGGCTTCAAGCTGAGTCCC 2520
Db 2461 CTGCACTGGGTGGAATTCAGAACAAAGTTCTACTCAGGCAAGGCTTCAAGCTGAGTCCC 2520
Qy 2521 TTTCACTTGGCTGCCACAGATGATAGGGCCACCTGAGGCTTCTGCGAGACCTTCTTCT 2580
Db 2521 TTTCACTTGGCTGCCACAGATGATAGGGCCACCTGAGGCTTCTGCGAGACCTTCTTCT 2580
Qy 2581 GACCTCTGAGGAGGAGGAGAAATAAGACGGTCCGCTTGGCAAAAAA 2640
Db 2581 GACCTCTGAGGAGGAGGAGAAATAAGACGGTCCGCTTGGCAAAAAA 2640

RESULT 3

ADR44118

ID ADR44118 standard; cDNA; 2640 BP.

XX ADR44118;

XX ADR44118;

DT 04-NOV-2004 (first entry)
XX Human 116-kDa osteoclast proton pump subunit (OC-116 kDa) protein cDNA.
DE
XX
XX Human; OC-116 kD; 116-kDa osteoclast proton pump subunit; therapy;
KW bone degradation; osteoporosis; osteoarthritis; ss; gene.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 58..2523
CDS
FT /*tag= a
FT /product= "Human 116-kDa osteoclast proton pump subunit
FT (OC-116 kDa) protein"
FT /partial
FT /note= "No stop codon"
XX
XX US6777537-B1.
XX
XX 17-AUG-2004.
XX
XX 18-JUL-2000; 2000US-00618304.
XX
XX 22-FEB-1996; 96US-00605378.
XX
XX (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.
XX
XX Stashenko P, Li Y;
XX
XX WPI; 2004-591304/57.
DR P-PSDB; ADR44119.
XX
XX New isolated human 116-kDa osteoclast (OC-116 kDa) proton pump subunit
PT polypeptide, useful for preventing and/or treating diseases with bone
PT degradation, such as osteoporosis and osteoarthritis.
XX
XX Claim 1; SEQ ID NO 1; 16pp; English.
XX
XX The present invention relates to a human 116-kDa osteoclast proton pump
CC subunit (OC-116 kDa) polypeptide and its encoding polynucleotide. The
CC invention is useful for the prevention and treatment of diseases or
CC conditions associated with aberrant expression or activity of the OC-116
CC kDa protein and bone degradation, such as osteoporosis and
CC osteoarthritis. The present sequence is human OC-116 kDa protein encoding
CC cDNA.
XX
XX Sequence 2640 BP; 445 A; 885 C; 824 G; 486 T; 0 U; 0 Other;
SQ

Query Match 100.0%; Score 2640; DB 13; Length 2640;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCGTGGCGGACGGGACGACGAGCGGCGGCGGCGAGCACACCCGGGACCATG 60
DB 1 CGGCGTGGCGGACGGGACGACGAGCGGCGGCGGCGAGCACACCCGGGACCATG 60

QY 61 GGCTCCATGTTCCGAGCGAGGAGTGGCCCTGGTCCAGCTCTTTCTGCCACAGCGGCT 120
DB 61 GGCTCCATGTTCCGAGCGAGGAGTGGCCCTGGTCCAGCTCTTTCTGCCACAGCGGCT 120

QY 121 GCCTACACCTGCTGAGTGGGCTGGGCGAGCTGGGCGCTCTGGAGTTTCAGAGCTCAAC 180
DB 121 GCCTACACCTGCTGAGTGGGCTGGGCGAGCTGGGCGCTCTGGAGTTTCAGAGCTCAAC 180

QY 181 GCCTCGGTGAGCGCTTCCAGAGACGCTTTGTGCTGATCTTTGGGCGCTGTGAGGAGCTG 240
DB 181 GCCTCGGTGAGCGCTTCCAGAGACGCTTTGTGCTGATCTTTGGGCGCTGTGAGGAGCTG 240

QY 241 GAGAAGACCTTCACTTCTTCGAGGAGGAGTGGCGGCTGGGCTGGTCTTCCGCCCG 300
DB 241 GAGAAGACCTTCACTTCTTCGAGGAGGAGTGGCGGCTGGGCTGGTCTTCCGCCCG 300

QY 301 CCAAAGGGGAGGCTGCGGCGACCCCCACCCCGGACCTGCTGCGATCCAGGAGGAGCG 360

Db 301 CCAAAGGGGAGGCTGCGGCGACCCCGGACCTGCTGCGATCCAGGAGGAGCG 360
QY 361 GAGCGCTGGCCAGGAGCTGCGGATGTGCGGGCAACAGCAGGCGCTTCCGGGCCAG 420
Db 361 GAGCGCTGGCCAGGAGCTGCGGATGTGCGGGCAACAGCAGGCGCTTCCGGGCCAG 420
QY 421 CTGCAACAGCTGAGCTTCAACGCGCGCTGTAGCGCAGGCGCATGAACCTCAGTGCA 480
Db 421 CTGCAACAGCTGAGCTTCAACGCGCGCTGTAGCGCAGGCGCATGAACCTCAGTGCA 480
QY 481 GCGGCCACACAGATGGGCTCAGAGAGACGCCCTGTCTCCAGGCGCCCGGGGGCGG 540
Db 481 GCGGCCACACAGATGGGCTCAGAGAGACGCCCTGTCTCCAGGCGCCCGGGGGCGG 540
QY 541 CACCAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGGTGGAGCCCGCACAGGCGCTGCC 600
Db 541 CACCAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGGTGGAGCCCGCACAGGCGCTGCC 600
QY 601 CTAGAGCGCTGCTGAGGCGCTGCGGCGCTTCTCTCATTTGCGAGCTTTCAGGAGCTG 660
Db 601 CTAGAGCGCTGCTGAGGCGCTGCGGCGCTTCTCTCATTTGCGAGCTTTCAGGAGCTG 660
QY 661 GAGCAGCGCTGGAGCACCCCGTGAGCGGCGAGCCAGCAGCTGGATGACCTTCTCATC 720
Db 661 GAGCAGCGCTGGAGCACCCCGTGAGCGGCGAGCCAGCAGCTGGATGACCTTCTCATC 720
QY 721 TCCTACTGCGGTGAGCAGATCGGACAGATCGGCAAGATCACGGACTGTCTTCACTGCG 780
Db 721 TCCTACTGCGGTGAGCAGATCGGACAGATCGGCAAGATCACGGACTGTCTTCACTGCG 780
QY 781 CAGCTCTTCCGTTTCTGACAGAGGAGGCGCGCTCGGGGCGCTGCGAGCAGCTGCA 840
Db 781 CAGCTCTTCCGTTTCTGACAGAGGAGGCGCGCTCGGGGCGCTGCGAGCAGCTGCA 840
QY 841 CAGCAGGCGCAGGAGCTGCGGAGGCTCTCGGGGAGCAGAGCGGTTCTTGAGCCAGTG 900
Db 841 CAGCAGGCGCAGGAGCTGCGGAGGCTCTCGGGGAGCAGAGCGGTTCTTGAGCCAGTG 900
QY 901 CTAGGCGCGGTGCTGACGCTGCTGCCCGCAGGCGAGGTGAGTCCACAGATGAAGGCC 960
Db 901 CTAGGCGCGGTGCTGACGCTGCTGCCCGCAGGCGAGGTGAGTCCACAGATGAAGGCC 960
QY 961 GTGTACTGCGCTGAAACAGTGCAGCTGAGCAGCACGACAGTGCCTCATTTGCCGAG 1020
Db 961 GTGTACTGCGCTGAAACAGTGCAGCTGAGCAGCACGACAGTGCCTCATTTGCCGAG 1020
QY 1021 GCCTGTGCTCTGTCGAGACCTGCGCGCTGCGAGGCGCTTCCGGGAGCTGCGATG 1080
Db 1021 GCCTGTGCTCTGTCGAGACCTGCGCGCTGCGAGGCGCTTCCGGGAGCTGCGATG 1080
QY 1081 GAGGAGGAGTGAAGTGCCTGCTCACCGCATCCCTGCGGGACATGCCCCCACAATC 1140
Db 1081 GAGGAGGAGTGAAGTGCCTGCTCACCGCATCCCTGCGGGACATGCCCCCACAATC 1140
QY 1141 ATCCGACCAACCGCTTCAAGCGCAGCTTCCAGGCGATCTGGATGCTACGCGTGGCG 1200
Db 1141 ATCCGACCAACCGCTTCAAGCGCAGCTTCCAGGCGATCTGGATGCTACGCGTGGCG 1200
QY 1201 CGCTACGAGAGTCAACCCCGCTTCCCTACACCATCATCACTTCCCTTCTTCTTGGCT 1260
Db 1201 CGCTACGAGAGTCAACCCCGCTTCCCTACACCATCATCACTTCCCTTCTTCTTGGCT 1260
QY 1261 GTGATGTTGGGAGTGTGGGCGACGGGCTGCTCATGTTCTTCTTGGCGCTGGCCATGCT 1320
Db 1261 GTGATGTTGGGAGTGTGGGCGACGGGCTGCTCATGTTCTTCTTGGCGCTGGCCATGCT 1320
QY 1321 CTTTGGGAGAACCGACCGGCTGTGAAGCGCGCAGAACGAGATCTGGCAGACTTCTTC 1380
Db 1321 CTTTGGGAGAACCGACCGGCTGTGAAGCGCGCAGAACGAGATCTGGCAGACTTCTTC 1380
QY 1381 AGGCGCGCTTACCTGCTCTGCTTATGGGCGCTTCTTCCATCTACACCGGCTTCATCTAC 1440

```
Db 1381 AGGGGCGGTACTCTGCTCTCTGCTTATGGGCGCTTCTCTCATCTACACCGGCTTCATCTAC 1440
QY 1441 AACAGTGTCTTCACTCGCGCCACACAGCATCTTCCCTCGGGCTGGAGTGTGGCGGCATG 1500
Db 1441 AACAGTGTCTTCACTCGCGCCACACAGCATCTTCCCTCGGGCTGGAGTGTGGCGGCATG 1500
QY 1501 GCCAACAGTCTGGCTGGAGTGTATGATATTCCTGGGCCAGACACAGATGCTTACCCCTGGAT 1560
Db 1501 GCCAACAGTCTGGCTGGAGTGTATGATATTCCTGGGCCAGACACAGATGCTTACCCCTGGAT 1560
QY 1561 CCAACAGTCAAGGCTGTCTTCTGGGACCTACCCCTTTGGGATCATCTTGGATCATTTGGAGC 1620
Db 1561 CCAACAGTCAAGGCTGTCTTCTGGGACCTACCCCTTTGGGATCATCTTGGGATCATTTGGAGC 1620
QY 1621 CTGGCTGCCAACACCTTGGAGCTTCTCAACTCTTCAAGATGAAGATGTCGGTCACTCTG 1680
Db 1621 CTGGCTGCCAACACCTTGGAGCTTCTCAACTCTTCAAGATGAAGATGTCGGTCACTCTG 1680
QY 1681 GGGCTCGTGCACATGCGCTTTGGGGTGGTCTCGGAGTCTTCAACACAGCTGCACCTTTGGC 1740
Db 1681 GGGCTCGTGCACATGCGCTTTGGGGTGGTCTCGGAGTCTTCAACACAGCTGCACCTTTGGC 1740
QY 1741 CAGAGCACCGGTGTCTGTGAGACGCTGCGGAGCTCACCTTCTGTGGAATCTTTC 1800
Db 1741 CAGAGCACCGGTGTCTGTGAGACGCTGCGGAGCTCACCTTCTGTGGAATCTTTC 1800
QY 1801 GGTACCTCGTGTCTAGTCTATCAAGTGGTGTGTGCTGGGCTGCCAGGGCGGC 1860
Db 1801 GGTACCTCGTGTCTAGTCTATCAAGTGGTGTGTGCTGGGCTGCCAGGGCGGC 1860
QY 1861 TCGCCAGCATCTCATCCACTTTCATCAACATGTTCTCTTCTCCACAGCCGCCAGCAAC 1920
Db 1861 TCGCCAGCATCTCATCCACTTTCATCAACATGTTCTCTTCTCCACAGCCGCCAGCAAC 1920
QY 1921 AGGTGCTCTACCCCGGAGAGGTGGTCCAGGCCACGCTGTGTGCTTGGCTTTGGCC 1980
Db 1921 AGGTGCTCTACCCCGGAGAGGTGGTCCAGGCCACGCTGTGTGCTTGGCTTTGGCC 1980
QY 1981 ATGGTGGCCATCTGTGCTGTGACACACCCCTGCACTCTGTCACCGCCAGCCGCGCCG 2040
Db 1981 ATGGTGGCCATCTGTGCTGTGACACACCCCTGCACTCTGTCACCGCCAGCCGCGCCG 2040
QY 2041 CTGGGAGGAGGCGCTGACCGACAGAGGAGAAACAAGGCGGGTGTGTGACCTTGCCT 2100
Db 2041 CTGGGAGGAGGCGCTGACCGACAGAGGAGAAACAAGGCGGGTGTGTGACCTTGCCT 2100
QY 2101 GAGCATCTGTGAATGGCTGGAGCTCCGATGAGGAGAAAGGAGGGGCTGTGATGATGAA 2160
Db 2101 GAGCATCTGTGAATGGCTGGAGCTCCGATGAGGAGAAAGGAGGGGCTGTGATGATGAA 2160
QY 2161 GAGGAGGCGAGCTGCTCCCTCCGAGGTGCTCATGCAACAGGCCATCCACACCATCGAG 2220
Db 2161 GAGGAGGCGAGCTGCTCCCTCCGAGGTGCTCATGCAACAGGCCATCCACACCATCGAG 2220
QY 2221 TTCTGCTGGCTGTGCTTCCAACACCGCTCTCTACTCTGCGCTGTGGGCGCTTGAGCCTG 2280
Db 2221 TTCTGCTGGCTGTGCTTCCAACACCGCTCTCTACTCTGCGCTGTGGGCGCTTGAGCCTG 2280
QY 2281 GCCACGCCAGTGTCCAGGTTCTGTGGGCTCATGGTGTATGGGCTATAGGCTTGGGCTG 2340
Db 2281 GCCACGCCAGTGTCCAGGTTCTGTGGGCTCATGGTGTATGGGCTATAGGCTTGGGCTG 2340
QY 2341 GSCCGGAGGTGGGCTGGGCTGTGGTGTGCTGCTCCCATCTTTGCGCCCTTGGCCG 2400
Db 2341 GSCCGGAGGTGGGCTGGGCTGTGGTGTGCTGCTCCCATCTTTGCGCCCTTGGCCG 2400
QY 2401 ATGACCGTGGCTATCTGCTGTGTGATGAGGAGTCTCTAGCCCTTCTGACGCGCTTGGCG 2460
Db 2401 ATGACCGTGGCTATCTGCTGTGTGATGAGGAGTCTCTAGCCCTTCTGACGCGCTTGGCG 2460
QY 2461 CTGCACTGGGTGGATTCAGAAACAAGTTCTACTAGGCAAGGCTTACAGCTGATGCC 2520
Db 2461 CTGCACTGGGTGGATTCAGAAACAAGTTCTACTAGGCAAGGCTTACAGCTGATGCC 2520
```

```
QY 2521 TTTCCTTTCGCTGCCACAGATGACTAGGGCCACTGAGGTCTTGCAGACCTCTTCT 2580
Db 2521 TTTCCTTTCGCTGCCACAGATGACTAGGGCCACTGAGGTCTTGCAGACCTCTTCT 2580
QY 2581 GACCTCTGAGGCGAGGAGGAATAAGACGGTCCGCCCTGGCAAAAAA 2640
Db 2581 GACCTCTGAGGCGAGGAGGAATAAGACGGTCCGCCCTGGCAAAAAA 2640

RESULT 4
ADY92605
ID ADY92605 standard; cDNA; 2640 BP.
XX
AC ADY92605;
XX
DT 02-JUN-2005 (first entry)
XX
DE Human 116 kD osteoclast proton pump OC-116 subunit cDNA.
XX
KW cloning; osteopathic; antiarthritic; gene therapy;
KW osteoclast proton pump; bone disease; osteoporosis; osteoarthritis; gene;
XX ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 58..2547
FT /tag= a
FT /product= "116 kD osteoclast proton pump subunit"
XX
XX US2005064448-A1.
XX
XX 24-MAR-2005.
XX
XX 20-FEB-2004; 2004US-00783519.
XX
XX 22-FEB-1996; 96US-00605378.
XX 18-JUL-2000; 2000US-00618304.
XX
XX (STAS/) STASHENKO P.
XX (LIYY/) LI Y.
XX
XX Stashenko P, Li Y;
XX
XX WPI; 2005-241271/25.
XX P-PSDB; ADY92606.
XX
XX New isolated genes and encoded human 116-kDa osteoclast proton pump
XX subunit polypeptides useful for treating bone mass disorders associated
XX with aberrant rate of bone degradation (e.g. osteoporosis or
XX osteoarthritis).
XX
XX Claim 3; SEQ ID NO 1; 17pp; English.
XX
XX The invention relates to an isolated gene encoding a polypeptide which is
XX a human 116-kD osteoclast proton pump subunit. The composition and
XX methods are useful for treating bone mass disorders characterized by
XX aberrant rate of bone degradation (e.g. osteoporosis or osteoarthritis).
XX This protein may also be used as cell surface markers for osteoclasts.
XX This sequence corresponds to the cDNA encoding the novel proton pump
XX subunit. The cDNA was isolated by differential screening of a human
XX osteoclastoma cDNA library.
XX
XX Sequence 2640 BP; 445 A; 885 C; 824 G; 486 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 2640; DB 14; Length 2640;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGCGTGGCGGACGCGGACGAGCGGCGGCGGACGAGCACACCGCGGACCATG 60
Db 1 CGGCGTGGCGGACGCGGACGAGCGGCGGCGGACGAGCACACCGCGGACCATG 60
```

QY 61 GGCCTCCATGTTCCGGAGCGAGAGGTGGCCCTGGTCCAGCTCTTTCTGCCCCACAGCGGCT 120
Db 61 GGCCTCCATGTTCCGGAGCGAGAGGTGGCCCTGGTCCAGCTCTTTCTGCCCCACAGCGGCT 120
QY 121 GCCTACACCTCGCTGAGTCGGCTGGCGAGCTGGGCCCTCGTGGAGTTCCAGAGACCTCAAC 180
Db 121 GCCTACACCTCGCTGAGTCGGCTGGCGAGCTGGGCCCTCGTGGAGTTCCAGAGACCTCAAC 180
QY 181 GCCTCGGTGAGCGCTTCCAGAGACGCTTTGTGTTGATGTTTGGCGCTGTGAGAGCTG 240
Db 181 GCCTCGGTGAGCGCTTCCAGAGACGCTTTGTGTTGATGTTTGGCGCTGTGAGAGCTG 240
QY 241 GAGAGACCTTCACTTCTTCCAGAGAGAGTGGCGGGCTGGGCTGGTCTCTGCCCGG 300
Db 241 GAGAGACCTTCACTTCTTCCAGAGAGAGTGGCGGGCTGGGCTGGTCTCTGCCCGG 300
QY 301 CCAAGAGGAGGCTGCGGGCACCCACCCCGGACCTGTGCGCATCCAGAGAGAGACG 360
Db 301 CCAAGAGGAGGCTGCGGGCACCCACCCCGGACCTGTGCGCATCCAGAGAGAGACG 360
QY 361 GAGCGCTGGCCAGAGAGCTGCGGGATGTGCGGGGCAACCAAGAGGCGCTGCGGGCCGAG 420
Db 361 GAGCGCTGGCCAGAGAGCTGCGGGATGTGCGGGGCAACCAAGAGGCGCTGCGGGCCGAG 420
QY 421 CTGCACAGCTGCAGCTCCAGCGCCGCTGTACGCCAGGGCCATGAACCTCAGCTGGCA 480
Db 421 CTGCACAGCTGCAGCTCCAGCGCCGCTGTACGCCAGGGCCATGAACCTCAGCTGGCA 480
QY 481 GCGGCCACACAGATGGGGCTCAGAGAGAGCGCCCTGTCTCCAGGCCCCCGGGGGCGG 540
Db 481 GCGGCCACACAGATGGGGCTCAGAGAGAGCGCCCTGTCTCCAGGCCCCCGGGGGCGG 540
QY 541 CACAGGACCTGAGGGTCAACTTTGTGGCAGGTGCCGTGGAGCCCCCAAGGCCCTTGC 600
Db 541 CACAGGACCTGAGGGTCAACTTTGTGGCAGGTGCCGTGGAGCCCCCAAGGCCCTTGC 600
QY 601 CTAGAGCGCTGCTCGAGGGCTGCGGGCTTCTCATTTGCGAGCTTCAGAGCTTCAGAGGCTG 660
Db 601 CTAGAGCGCTGCTCGAGGGCTGCGGGCTTCTCATTTGCGAGCTTCAGAGCTTCAGAGGCTG 660
QY 661 GAGCAGCGCTGGAGCACCCCGTCACCGGGGAGCGAGCCAGCTGGATGACCTTCTCATC 720
Db 661 GAGCAGCGCTGGAGCACCCCGTCACCGGGGAGCGAGCCAGCTGGATGACCTTCTCATC 720
QY 721 TCCTACTGGGGTGAAGAGATCGGACAGAGATCCGCAAGATCAAGGACTGCTTCCACTGC 780
Db 721 TCCTACTGGGGTGAAGAGATCGGACAGAGATCCGCAAGATCAAGGACTGCTTCCACTGC 780
QY 781 CAGCTCTTCCGGTTTCTGACAGGAGAGAGGCGCCGCTCGGGCCCTGCGAGCAGCTGCAA 840
Db 781 CAGCTCTTCCGGTTTCTGACAGGAGAGAGGCGCCGCTCGGGCCCTGCGAGCAGCTGCAA 840
QY 841 CAGCAGAGCAGGAGCTGCAAGAGGTCTCGGGGAGACAGAGCGGTTCTTGGAGCAGGTG 900
Db 841 CAGCAGAGCAGGAGCTGCAAGAGGTCTCGGGGAGACAGAGCGGTTCTTGGAGCAGGTG 900
QY 901 CTAGGCGGGTGTGCAAGCTGCTGCGCCAGGGCAGGTGCAAGATGAAGGCC 960
Db 901 CTAGGCGGGTGTGCAAGCTGCTGCGCCAGGGCAGGTGCAAGATGAAGGCC 960
QY 961 GTGTACTTGGCCCTGAACCAAGTCAGAGTCAGAGCAACAGCAAGTGCCTCATTTGCCGAG 1020
Db 961 GTGTACTTGGCCCTGAACCAAGTCAGAGTCAGAGCAACAGCAAGTGCCTCATTTGCCGAG 1020
QY 1021 GCCTGTGTCTGTGCGAGACCTGCCCGCTCGAGGAGCGCTGCGGAGCAGCTCGATG 1080
Db 1021 GCCTGTGTCTGTGCGAGACCTGCCCGCTCGAGGAGCGCTGCGGAGCAGCTCGATG 1080
QY 1081 GAGGAGGAGTGTGCTGGCTCACCGCATCCCTGCGGGACATGCCCCCACCACCTC 1140
Db 1081 GAGGAGGAGTGTGCTGGCTCACCGCATCCCTGCGGGACATGCCCCCACCACCTC 1140

QY 1141 ATCCGACCAACCCGCTTCAAGGCGAGCTTCCAGGGCATCGTGGATCGCTACGCGCTGGGC 1200
Db 1141 ATCCGCGCAACCAACCCGCTTCAAGGCGAGCTTCCAGGGCATCGTGGATCGCTACGCGCTGGGC 1200
QY 1201 CGCTACAGGAGGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCTCTGCTTTGCT 1260
Db 1201 CGCTACAGGAGGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCTCTGCTTTGCT 1260
QY 1261 GTGATGTTGCGGGATGTGGCCACCGGCTGCTCATGTTCTTCTTCTGCGCTTGGCCATGGTC 1320
Db 1261 GTGATGTTGCGGGATGTGGCCACCGGCTGCTCATGTTCTTCTTCTGCGCTTGGCCATGGTC 1320
QY 1321 CTTGCGGAGAACCCGACCGGCTGTGAAGCCGCGCAGAACAGATCTGCGAGACTTCTTC 1380
Db 1321 CTTGCGGAGAACCCGACCGGCTGTGAAGCCGCGCAGAACAGATCTGCGAGACTTCTTC 1380
QY 1381 AGGGGCGCTACCTGCTCTCTGCTTATGGGCTGTCTTCTCATCTACACCGGCTTCACTAC 1440
Db 1381 AGGGGCGCTACCTGCTCTCTGCTTATGGGCTGTCTTCTCATCTACACCGGCTTCACTAC 1440
QY 1441 AACGAGTCTTCACTGCGGCCACCGAGCATTTTCCCTCTGGGCTGGAGTGTGGCCCATG 1500
Db 1441 AACGAGTCTTCACTGCGGCCACCGAGCATTTTCCCTCTGGGCTGGAGTGTGGCCCATG 1500
QY 1501 GCCAACAGCTTGGCTGAGTGTGATTCCTGCGCCACGACACAGATGCTTACCTGAT 1560
Db 1501 GCCAACAGCTTGGCTGAGTGTGATTCCTGCGCCACGACACAGATGCTTACCTGAT 1560
QY 1561 CCCAAAGCTACCGGTGTCTTCTGGGACCTTACCCCTTTGGCATCGATCTTATTTGGAGC 1620
Db 1561 CCCAAAGCTACCGGTGTCTTCTGGGACCTTACCCCTTTGGCATCGATCTTATTTGGAGC 1620
QY 1621 CTGCTGCCAACCACTTGAAGTCTTCAAGTGAAGATGTCGCTCATCTG 1680
Db 1621 CTGCTGCCAACCACTTGAAGTCTTCAAGTGAAGATGTCGCTCATCTG 1680
QY 1681 GCGCTGTCGACATGGCTTTGGGGTGGTCTCGAGTCTTCAACAGCTGCTTGGC 1740
Db 1681 GCGCTGTCGACATGGCTTTGGGGTGGTCTCGAGTCTTCAACAGCTGCTTGGC 1740
QY 1741 CAGAGGACCGGCTGCTGGAGCGCTGCGGAGCTCACTTCTTCTGCTGGACTCTTC 1800
Db 1741 CAGAGGACCGGCTGCTGGAGCGCTGCGGAGCTCACTTCTTCTGCTGGACTCTTC 1800
QY 1801 GGTTCCTGCTGTTCTAGTCACTACAAGTGGTGTGTGCTGGGCTGCGAGGCGCGCC 1860
Db 1801 GGTTCCTGCTGTTCTAGTCACTACAAGTGGTGTGTGCTGGGCTGCGAGGCGCGCC 1860
QY 1861 TCGCCAGCATCTCATCTTCAACATGTTCTCTTCTTCTCCACAGCCCCAGCAAC 1920
Db 1861 TCGCCAGCATCTCATCTTCAACATGTTCTCTTCTTCTCCACAGCCCCAGCAAC 1920
QY 1921 AGGCTGCTTACCCCGGAGGAGTGTCCAGGCGACGCTGCTGCTGCTGGCTTGGCC 1980
Db 1921 AGGCTGCTTACCCCGGAGGAGTGTCCAGGCGACGCTGCTGCTGCTGGCTTGGCC 1980
QY 1981 ATGCTGCCATCTCTGCTTGGCAACCCCTGCACTGCTGCACTGCGCCAGCGCGCCG 2040
Db 1981 ATGCTGCCATCTCTGCTTGGCAACCCCTGCACTGCTGCACTGCGCCAGCGCGCCG 2040
QY 2041 CTTGCGAGGAGGCGCTGACCCGACAGGAGAAACAAAGGCGGGTGTGCTGGACTGCT 2100
Db 2041 CTTGCGAGGAGGCGCTGACCCGACAGGAGAAACAAAGGCGGGTGTGCTGGACTGCT 2100
QY 2101 GAGCATCTGTGAATGCTGAGCTCCGATGAGGAAAGGAGGAGGCTGGATGATGAA 2160
Db 2101 GAGCATCTGTGAATGCTGAGCTCCGATGAGGAAAGGAGGAGGCTGGATGATGAA 2160
QY 2161 GAGGAGGCGAGCTGCTCCCTCCGAGGTGCTCATGCAACAGGCGCATCCACACCATCGAG 2220
Db 2161 GAGGAGGCGAGCTGCTCCCTCCGAGGTGCTCATGCAACAGGCGCATCCACACCATCGAG 2220
QY 2221 TTCTGCTGCGGCTGCGTCTCAAACACCGCTCTCACTGCGGCTGTGGGCGCTGAGCGCTG 2280

```
Db 2221 TTCTGCTGGGCTGGCTCTCCACACCGCTCTCTACTGCGCTGTGGGCTCTGAGCCTG 2280
Qy 2281 GCCACGCCAGCTGTCCAGGTTCTGTGGGCTCATGTTGATGCGCATAGGCTTGGGCTG 2340
Db 2281 GCCACGCCAGCTGTCCAGGTTCTGTGGGCTCATGTTGATGCGCATAGGCTTGGGCTG 2340
Qy 2341 GGCCGGAGGTGGGCTGGGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 2400
Db 2341 GGCCGGAGGTGGGCTGGGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 2400
Qy 2401 ATGACCGTGGCTATCTCTGCTGTGGTGTGGGAGGACTCTCTGACGCTTCTCTGACGCTTGGCGG 2460
Db 2401 ATGACCGTGGCTATCTCTGCTGTGGTGTGGGAGGACTCTCTGACGCTTCTCTGACGCTTGGCGG 2460
Qy 2461 CTGCACTGGGTGAATTCAGAAACAAGTTCTACTCAGGCAAGGCTTACAAAGTGTGATCCC 2520
Db 2461 CTGCACTGGGTGAATTCAGAAACAAGTTCTACTCAGGCAAGGCTTACAAAGTGTGATCCC 2520
Qy 2521 TTCACTTCTGCTCCACAGATGACTAGGCGCCACTGAGTCTGCTGCGAGCTTCTTCT 2580
Db 2521 TTCACTTCTGCTCCACAGATGACTAGGCGCCACTGAGTCTGCTGCGAGCTTCTTCT 2580
Qy 2581 GACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
Db 2581 GACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
```

RESULT 5

```
ABL64400
ID ABL64400 standard; DNA; 2655 BP.
XX
AC ABL64400;
XX
DT 15-MAY-2002 (first entry)
XX
DE Stomach cancer related gene sequence SEQ ID NO:2737.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; db.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
```

```
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
```

```
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 29-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
PS Claim 1; SEQ ID NO 2737; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
SQ Sequence 2655 BP; 460 A; 885 C; 824 G; 486 T; 0 U; 0 Other;
Query Match 100.0%; Score 2640; DB 6; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGGCGTGGCGGACCGGCGGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
Db 1 CGGCGTGGCGGACCGGCGGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
Qy 61 GGCTCCATGTTCCCGAGCGGAGGAGTGGCCCTGGTCCAGCTCTTTTCCCGCCACAGCGGCT 120
Db 61 GGCTCCATGTTCCCGAGCGGAGGAGTGGCCCTGGTCCAGCTCTTTTCCCGCCACAGCGGCT 120
Qy 121 GGCTACACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180
Db 121 GGCTACACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 180
Qy 181 GGCTCGGTGAGCGCTTCCAGAGACGCTTTGTGGTGTGATGTTGTGGCGCTGTGAGGAGCTG 240
Db 181 GGCTCGGTGAGCGCTTCCAGAGACGCTTTGTGGTGTGATGTTGTGGCGCTGTGAGGAGCTG 240
```

QY 241 GAGAGACCTTCACTTCTGACAGAGAGAGTGGCGGGGCTGGGCTGGTCTGCGCCCG 300
DB 241 GAGAGACCTTCACTTCTGACAGAGAGAGTGGCGGGGCTGGGCTGGTCTGCGCCCG 300
QY 301 CCAAAGGGAGGCTGCGGCACCCCAACCCCGGACCTGCTGCGCATCCAGAGGAGAGC 360
DB 301 CCAAAGGGAGGCTGCGGCACCCCAACCCCGGACCTGCTGCGCATCCAGAGGAGAGC 360
QY 361 GAGCGCTGCGCCAGGAGCTGCGGATGTGCGGGCAACCAAGAGGCCCTGCGGGCCAG 420
DB 361 GAGCGCTGCGCCAGGAGCTGCGGATGTGCGGGCAACCAAGAGGCCCTGCGGGCCAG 420
QY 421 CTGACACAGCTGACGCTCCAGCGCGCGTGTACGCCAGGGCCATGAACCTCAGCTGGCA 480
DB 421 CTGACACAGCTGACGCTCCAGCGCGCGTGTACGCCAGGGCCATGAACCTCAGCTGGCA 480
QY 481 GCGGCCCAACAGATGGGGCTCAGAGAGGAGCGCCCTGCTCCAGGGCCCCCGGGGGCG 540
DB 481 GCGGCCCAACAGATGGGGCTCAGAGAGGAGCGCCCTGCTCCAGGGCCCCCGGGGGCG 540
QY 541 CACAGGACCTGAGGGTCACTTTGTGGCAGGTGCGGTGGAGCCCCCAAGGCCCTGCC 600
DB 541 CACAGGACCTGAGGGTCACTTTGTGGCAGGTGCGGTGGAGCCCCCAAGGCCCTGCC 600
QY 601 CTAGAGCGCTGCTGAGAGGCTGCGGGCTTCTCATTTGCGAGCTTCAGGGAGCTG 660
DB 601 CTAGAGCGCTGCTGAGAGGCTGCGGGCTTCTCATTTGCGAGCTTCAGGGAGCTG 660
QY 661 GAGCAGCGCTGGAGCACCCGTCGACGGGGAGCGACGACGTCGATGACCTTCTCATC 720
DB 661 GAGCAGCGCTGGAGCACCCGTCGACGGGGAGCGACGACGTCGATGACCTTCTCATC 720
QY 721 TCCTACTGGGGTGAAGAGATCGGACAGAAGATCCGCAAGATCAAGGACTGCTTCCACTGC 780
DB 721 TCCTACTGGGGTGAAGAGATCGGACAGAAGATCCGCAAGATCAAGGACTGCTTCCACTGC 780
QY 781 CAGCTTTCCTGTTCTGACAGAGAGAGGCGCGCTCGGGGCTTGCAGAGCTGCAAG 840
DB 781 CAGCTTTCCTGTTCTGACAGAGAGAGGCGCGCTCGGGGCTTGCAGAGCTGCAAG 840
QY 841 CAGCAGAGCAGAGCTGACAGAGTCTCGGGGAGACAGAGCGGTTCTGAGCAGGTG 900
DB 841 CAGCAGAGCAGAGCTGACAGAGTCTCGGGGAGACAGAGCGGTTCTGAGCAGGTG 900
QY 901 CTAGGCGGGTGTGACGCTGCTGCGCCAGGGAGGTGACAGTCCCAAGATGAAGGCC 960
DB 901 CTAGGCGGGTGTGACGCTGCTGCGCCAGGGAGGTGACAGTCCCAAGATGAAGGCC 960
QY 961 GTGTACTGGCCCTGAACCAAGTGCAGGTGAGACCAAGCAAGTGCCTCATTTGCCGAG 1020
DB 961 GTGTACTGGCCCTGAACCAAGTGCAGGTGAGACCAAGCAAGTGCCTCATTTGCCGAG 1020
QY 1021 GCCTGGTGTCTGTGCGAGACCTGCGCGCTGACAGAGGCCCTGCGGGACAGCTCGATG 1080
DB 1021 GCCTGGTGTCTGTGCGAGACCTGCGCGCTGACAGAGGCCCTGCGGGACAGCTCGATG 1080
QY 1081 GAGGAGGAGTGAAGTGCCTGGTGTGACCGCATCCCTGCGGGACATGCCCCCACACTC 1140
DB 1081 GAGGAGGAGTGAAGTGCCTGGTGTGACCGCATCCCTGCGGGACATGCCCCCACACTC 1140
QY 1141 ATCCGACCAACCGCTTACCGGCGAGTTCCAGGGCATCGTGGATCGTACGGGCTGGGC 1200
DB 1141 ATCCGACCAACCGCTTACCGGCGAGTTCCAGGGCATCGTGGATCGTACGGGCTGGGC 1200
QY 1201 CGCTACAGAGGTCAACCCCGCTCCCTACACCATCATCCTTCCCTTCTGTTGCT 1260
DB 1201 CGCTACAGAGGTCAACCCCGCTCCCTACACCATCATCCTTCCCTTCTGTTGCT 1260
QY 1261 GTGATGTTGGGGATGTGGGCCACGGGCTGCTCATGTTCTTTCGCGCTGGGCATGGTC 1320
DB 1261 GTGATGTTGGGGATGTGGGCCACGGGCTGCTCATGTTCTTTCGCGCTGGGCATGGTC 1320

QY 1321 CTTGCGGAGAACCGACCGGCTGTGAAGCCCGCAGAACAGATCTGCGAGACTTCTTTC 1380
DB 1321 CTTGCGGAGAACCGACCGGCTGTGAAGCCCGCAGAACAGATCTGCGAGACTTCTTTC 1380
QY 1381 AGGGGCGCTACCTGCTTCTGCTTATGGGCTGTTCCTCATCTACACGGCTTCACTAC 1440
DB 1381 AGGGGCGCTACCTGCTTCTGCTTATGGGCTGTTCCTCATCTACACGGCTTCACTAC 1440
QY 1441 AACGAGTGTCTTCACTGCGGCCACAGCATCTTCCCTCGGGCTGGAGTGTGGCGCCATG 1500
DB 1441 AACGAGTGTCTTCACTGCGGCCACAGCATCTTCCCTCGGGCTGGAGTGTGGCGCCATG 1500
QY 1501 GCCAACAGTGTGCTGAGTGTGATGCTTCTGCGGCCAGCACAGATGCTTACCTGAT 1560
DB 1501 GCCAACAGTGTGCTGAGTGTGATGCTTCTGCGGCCAGCACAGATGCTTACCTGAT 1560
QY 1561 CCCAACGTCACCGGTGTCTTCTGGGACCTTACCCCTTGGCATCGATCTTATTTGGAGC 1620
DB 1561 CCCAACGTCACCGGTGTCTTCTGGGACCTTACCCCTTGGCATCGATCTTATTTGGAGC 1620
QY 1621 CTGCTGCCAACCACTTGAAGTCTTCAAGTGTTCCTTCAAGATGAAGATGTCCGTCATCCTG 1680
DB 1621 CTGCTGCCAACCACTTGAAGTCTTCAAGTGTTCCTTCAAGATGAAGATGTCCGTCATCCTG 1680
QY 1681 GGCGTCTGACATGGGCTTGGGCTGCTCGAGTCTTCAACAGAGTGCATTTGGC 1740
DB 1681 GGCGTCTGACATGGGCTTGGGCTGCTCGAGTCTTCAACAGAGTGCATTTGGC 1740
QY 1741 CAGAGGACCGGCTGCTGGAGACCTGCGGAGCTGCTTCAACAGAGTGCATTTGGC 1800
DB 1741 CAGAGGACCGGCTGCTGGAGACCTGCGGAGCTGCTTCAACAGAGTGCATTTGGC 1800
QY 1801 GGTGTACCTGCTTCTAGTCACTACAAGTGGTGTGTCTGGGCTGCCAGGCGCCGCC 1860
DB 1801 GGTGTACCTGCTTCTAGTCACTACAAGTGGTGTGTCTGGGCTGCCAGGCGCCGCC 1860
QY 1861 TCGCCAGCATCTCATCTTCAATCAATGTTCTTCTTCTCCACAGGCGCCAGCAAC 1920
DB 1861 TCGCCAGCATCTCATCTTCAATCAATGTTCTTCTTCTCCACAGGCGCCAGCAAC 1920
QY 1921 AGGCTGCTTACCCCGGAGGAGTGTCCAGGACCTGCTGAGTGTCTTCTTCTCCAGGCGCC 1980
DB 1921 AGGCTGCTTACCCCGGAGGAGTGTCCAGGACCTGCTGAGTGTCTTCTTCTCCAGGCGCC 1980
QY 1981 ATGCTGCCATCTGCTTGTGGCACACCTTGCACCTGCTGCAACCGCACCGCGCGCC 2040
DB 1981 ATGCTGCCATCTGCTTGTGGCACACCTTGCACACCTTGCACACCGCGCGCC 2040
QY 2041 CTGGGAGGAGGCGCTGACCGACAGAGAGAAACAAAGGCGGGTGTCTGGACCTGCT 2100
DB 2041 CTGGGAGGAGGCGCTGACCGACAGAGAGAAACAAAGGCGGGTGTCTGGACCTGCT 2100
QY 2101 GAGCATCTCTGAATGCTGAGCTCCGATGAGAAAGAGGCGGGCTGGATGATGA 2160
DB 2101 GAGCATCTCTGAATGCTGAGCTCCGATGAGAAAGAGGCGGGCTGGATGATGA 2160
QY 2161 GAGGAGGCGGAGCTGCTCCCTCGAGTGTCTCATGACAGGCGCATCCACCATCGAG 2220
DB 2161 GAGGAGGCGGAGCTGCTCCCTCGAGTGTCTCATGACAGGCGCATCCACCATCGAG 2220
QY 2221 TTCTGCTGGGCTGCGTCTCAAACACCGCTTCTTACCTGCGGCTGTGGGCTGTGAGCTG 2280
DB 2221 TTCTGCTGGGCTGCGTCTCAAACACCGCTTCTTACCTGCGGCTGTGGGCTGTGAGCTG 2280
QY 2281 GCCACGCCAGCTGCTCCGAGGTTCTGTGGGCAATGTGTGATGAGGCTGGGCTG 2340
DB 2281 GCCACGCCAGCTGCTCCGAGGTTCTGTGGGCAATGTGTGATGAGGCTGGGCTG 2340
QY 2341 GGCGGAGGCTGGGCTGTGGGCTGTGGTGTCTTGTCCCATCTTTGCGGCTTTGCGGCTG 2400
DB 2341 GGCGGAGGCTGGGCTGTGGGCTGTGGTGTCTTGTCCCATCTTTGCGGCTTTGCGGCTG 2400
QY 2401 ATGACGCTGCTATCTGCTGGTGTGAGGAGTCTCAGGCTTCTGACGCTTCTGCGGCTGCGG 2460

| | | | |
|----|------|--|------|
| Qy | 421 | CTGCACCAAGCTTGCAGCTTCCACGCCCGCGTGTCTACGCCAGGGCCATGAACCTCAGCTGGCA | 480 |
| Db | 421 | | |
| Qy | 421 | CTGCACCAAGCTTGCAGCTTCCACGCCCGCGTGTCTACGCCAGGGCCATGAACCTCAGCTGGCA | 480 |
| Db | 421 | | |
| Qy | 481 | GCCGCCCACACAGATGGGGCTTCAGAGAGACGCGCCCTGTCTCAGGCGCCCGGGGGGCGG | 540 |
| Db | 481 | | |
| Qy | 541 | CACAGAGACTTGAGGGTCAACTTTTGTGGAGGTGCGGTGGAGAGCCCAAGAGCCCTTGC | 600 |
| Db | 541 | | |
| Qy | 601 | CTAGAGCCCTGTCTCTGAGGGCTGCGCGGCTTCTCATTTGCCAGCTTCAGGAGCTG | 660 |
| Db | 601 | | |
| Qy | 661 | GAGCAGCCGCTGGAGCACCCCGTGCAGGGCGAGCCAGCCACGTGGATGACCTTCTCATC | 720 |
| Db | 661 | | |
| Qy | 721 | TCCTACTGGGTGAGCAGATCGGACAGAAGATTCGCAAGATCAAGATCTCCACTGC | 780 |
| Db | 721 | | |
| Qy | 781 | CAGCTCTTCCGTTTCTGCACGAGGAGAGCGCCGCTCGGGGCCCTGCACAGCTGCA | 840 |
| Db | 781 | | |
| Qy | 841 | CAGCAGAGCCAGGAGCTTCAGGAGGTCTCTCGGGGAGACAGAGCGGTTCTTGAGCCAGGTG | 900 |
| Db | 841 | | |
| Qy | 901 | CTAGCCGGGTGTGCAGCTGCTCGCCACGAGGAGGTGCAGGTCCACAAGATGAAGGCC | 960 |
| Db | 901 | | |
| Qy | 961 | GTGTACCTGGCCCTGAAACAGTGAGCTGAGCACCAAGCAGACAGTGCCTCATTTGCCGAG | 1020 |
| Db | 961 | | |
| Qy | 1021 | GCCTGGTCTCTGTGCGAGACCTTGCCCGCTTCAGGAGGCGCTTCGCGGACAGCTCGATG | 1080 |
| Db | 1021 | | |
| Qy | 1081 | GAGAGGAGTGAGTGGCTGGCTACCGCATTCCTGTCGCGGACATGCCCCCACACTC | 1140 |
| Db | 1081 | | |
| Qy | 1141 | ATCCGCACCAACCGCTTCACGGCCAGCTTCAGGGCATCGTGGATCGCTACGGCGTGGC | 1200 |
| Db | 1141 | | |
| Qy | 1201 | CGCTACCAAGGAGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCTCTTTGCT | 1260 |
| Db | 1201 | | |
| Qy | 1261 | GTGATGTTGCGGATGTGGGCCACAGGGTGCTCATGTTCTCTTTGCGCCATGGCCATGGTC | 1320 |
| Db | 1261 | | |
| Qy | 1321 | CTTCGGGAGAACCAACCGGCTGTGAAAGCCGCGAGAACGAGATCTGGCAGACTTTCTTC | 1380 |
| Db | 1321 | | |
| Qy | 1381 | AGGGCCGCTACCTGCTCTCTTATGGGCTGTTCCTCATCTACACCGGCTTCATCTAC | 1440 |
| Db | 1381 | | |
| Qy | 1441 | AACGAGTGCTTCAGTTCGCGCACAGCATTTCCCTTCGGGCTGGAGTGTGGCCCATG | 1500 |
| Db | 1441 | | |

| | | | |
|----|------|---|------|
| QY | 1501 | GCACACAGTCTGGCTGGATGATGCAATTCCTGGCCAGGACACAGATGCTTACCTCTGGAT | 1560 |
| DB | 1501 | GCACACAGTCTGGCTGGATGATGCAATTCCTGGCCAGGACACAGATGCTTACCTCTGGAT | 1560 |
| QY | 1561 | CCCAACGTCAACGGTGTCCTTCTGGGACACCTACCCCTTTGGCATCGATCCTATTTGGAGC | 1620 |
| DB | 1561 | CCCAACGTCAACGGTGTCCTTCTGGGACACCTACCCCTTTGGCATCGATCCTATTTGGAGC | 1620 |
| QY | 1621 | CTGGCTGCCAACCACTTTGAGCTTCTCAACTCTCTTCAAGATGAAGATGTCCTGATCTCTG | 1680 |
| DB | 1621 | CTGGCTGCCAACCACTTTGAGCTTCTCAACTCTCTTCAAGATGAAGATGTCCTGATCTCTG | 1680 |
| QY | 1681 | GGCGTCTGCACATGGCGTTTGGGGTGTCCTCGAGTCTTCAACACACGTGCACTTTGGC | 1740 |
| DB | 1681 | GGCGTCTGCACATGGCGTTTGGGGTGTCCTCGAGTCTTCAACACACGTGCACTTTGGC | 1740 |
| QY | 1741 | CAGAGGCACCGGTCTGCTGGAGACGCTGCGGAGCTCACCTTCTGCTGGAGCTCTTTC | 1800 |
| DB | 1741 | CAGAGGCACCGGTCTGCTGGAGACGCTGCGGAGCTCACCTTCTGCTGGAGCTCTTTC | 1800 |
| QY | 1801 | GGTTACCTCGTGTTCCTAGTCACTACAAGTGGCTGTGTCTTGGGCTGCCAGGCGCGCC | 1860 |
| DB | 1801 | GGTTACCTCGTGTTCCTAGTCACTACAAGTGGCTGTGTCTTGGGCTGCCAGGCGCGCC | 1860 |
| QY | 1861 | TCGCCCCAGCATCTCATCTTCAATCATGTTTCTTCTTCCACAGCCCCAGCAAC | 1920 |
| DB | 1861 | TCGCCCCAGCATCTCATCTTCAATCATGTTTCTTCTTCCACAGCCCCAGCAAC | 1920 |
| QY | 1921 | AGGCTGCTTACCCCGGACAGGAGTGTCAGGCCACGCTGGTGGTCTGTCGCTTGGCC | 1980 |
| DB | 1921 | AGGCTGCTTACCCCGGACAGGAGTGTCAGGCCACGCTGGTGGTCTGTCGCTTGGCC | 1980 |
| QY | 1981 | ATGCTGCCCATCTCTGCTTGGCACACCCCTGCACCTCTGTCACCGCCACCGCGCGCCG | 2040 |
| DB | 1981 | ATGCTGCCCATCTCTGCTTGGCACACCCCTGCACCTCTGTCACCGCCACCGCGCGCCG | 2040 |
| QY | 2041 | CTGCGGAGGAGGCCGTGACCGACAGGAGGAAAAAAGGCGCGGGTTGTGCACTTGCCT | 2100 |
| DB | 2041 | CTGCGGAGGAGGCCGTGACCGACAGGAGGAAAAAAGGCGCGGGTTGTGCACTTGCCT | 2100 |
| QY | 2101 | GAGCATCTGTGAATGGCTGGAGTCCGATGAGGAAAAAGGCGGGGCTGTGATGAA | 2160 |
| DB | 2101 | GAGCATCTGTGAATGGCTGGAGTCCGATGAGGAAAAAGGCGGGGCTGTGATGAA | 2160 |
| QY | 2161 | GAGGAGGCCAGTCTGTCCTCCAGAGTGCTCATGACACAGGCCCATCCACACCATCGAG | 2220 |
| DB | 2161 | GAGGAGGCCAGTCTGTCCTCCAGAGTGCTCATGACACAGGCCCATCCACACCATCGAG | 2220 |
| QY | 2221 | TTCTGCTGGGCTGCGTCTCCACACCGCTCTCTACTGCGCTGTGGGCGCTGAGCCTG | 2280 |
| DB | 2221 | TTCTGCTGGGCTGCGTCTCCACACCGCTCTCTACTGCGCTGTGGGCGCTGAGCCTG | 2280 |
| QY | 2281 | GCCACGCCAGCTGTCCGAGTTCTGTGGGCGATGGTGATCGCATAGGCCCTGGGCTG | 2340 |
| DB | 2281 | GCCACGCCAGCTGTCCGAGTTCTGTGGGCGATGGTGATCGCATAGGCCCTGGGCTG | 2340 |
| QY | 2341 | GGCGGAGGTGGCGGTGGCGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 2400 |
| DB | 2341 | GGCGGAGGTGGCGGTGGCGGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG | 2400 |
| QY | 2401 | ATGACCGTGGCTATCTGCTGGTGATGAGGAGCTCTCAGCCTTCTGCAAGCCCTGGGG | 2460 |
| DB | 2401 | ATGACCGTGGCTATCTGCTGGTGATGAGGAGCTCTCAGCCTTCTGCAAGCCCTGGGG | 2460 |
| QY | 2461 | CTGCACTGGGTGAATTCAGAAAGATTTCTACTAGGCAAGGCTTACAGCTGAGTCCC | 2520 |
| DB | 2461 | CTGCACTGGGTGAATTCAGAAAGATTTCTACTAGGCAAGGCTTACAGCTGAGTCCC | 2520 |
| QY | 2521 | TTCACTTCTGCTGCCACAGATGAGTGGGCCCACTGCAAGTCTCTGCAAGCTTCTCT | 2580 |
| DB | 2521 | TTCACTTCTGCTGCCACAGATGAGTGGGCCCACTGCAAGTCTCTGCAAGCTTCTCT | 2580 |
| QY | 2581 | GACCTCTGAGGACAGGAGGAATAAGACGGTCCGCGCTGGCAAAAAAATAAAAAA | 2640 |

Qy 901 CTAGCCGGGTGCTGCAGCTGCTGCGCCAGGGCAGGTGCGAGTCCACAAGATGAAGGCC 960
Db 901 CTAGCCGGGTGCTGCAGCTGCTGCGCCAGGGCAGGTGCGAGTCCACAAGATGAAGGCC 960
Qy 961 GTGTACCTGGCCCTGAAACAGTGCAGGTGAGCACCAGCAAGTGCCTCATTTGCCGAG 1020
Db 961 GTGTACCTGGCCCTGAAACAGTGCAGGTGAGCACCAGCAAGTGCCTCATTTGCCGAG 1020
Qy 1021 GCCTGGTGTCTGTGCGAGACTCCCGCCCTGCGAGGGCCCTGCGGAGCAGCTCGATG 1080
Db 1021 GCCTGGTGTCTGTGCGAGACTCCCGCCCTGCGAGGGCCCTGCGGAGCAGCTCGATG 1080
Qy 1081 GAGGAGGAGTGAAGTGCCTGCTCAGCGCATCCCTGCGGAGATGCCGCCACACATC 1140
Db 1081 GAGGAGGAGTGAAGTGCCTGCTCAGCGCATCCCTGCGGAGATGCCGCCACACATC 1140
Qy 1141 ATCCGCAACCAACCGCTTTCAGCGCCAGCTTCCAGGGCATCGTGGATCGCTACGGGTGGGC 1200
Db 1141 ATCCGCAACCAACCGCTTTCAGCGCCAGCTTCCAGGGCATCGTGGATCGCTACGGGTGGGC 1200
Qy 1201 CGCTACAGGAGGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCTCTGTTGCT 1260
Db 1201 CGCTACAGGAGGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCTCTGTTGCT 1260
Qy 1261 GTGATGTTGGGGATGTGGGCCACGGGCTGCTCATGTTCTTCTTCCGCTGGCCATGGTC 1320
Db 1261 GTGATGTTGGGGATGTGGGCCACGGGCTGCTCATGTTCTTCTTCCGCTGGCCATGGTC 1320
Qy 1321 CTTGCGGAGAACCGACCGGTGTCAAGCCGCGCAGAACGAGATCTGCGAGACTTCTTC 1380
Db 1321 CTTGCGGAGAACCGACCGGTGTCAAGCCGCGCAGAACGAGATCTGCGAGACTTCTTC 1380
Qy 1381 AGGGCCGCTACCTGCTCTGCTTATGGGCTGTCTCCATCTACACCGGCTTCACTAC 1440
Db 1381 AGGGCCGCTACCTGCTCTGCTTATGGGCTGTCTCCATCTACACCGGCTTCACTAC 1440
Qy 1441 AACGAGTCTTACGTCGGCCACAGCATCTTCCCTCGGGCTGGAGTGGCGCCGATG 1500
Db 1441 AACGAGTCTTACGTCGGCCACAGCATCTTCCCTCGGGCTGGAGTGGCGCCGATG 1500
Qy 1501 GCCAACAGTCTGCTGCGAGTGTGATCTCTGCGCCAGCACAGATGCTTACCCCTGGAT 1560
Db 1501 GCCAACAGTCTGCTGCGAGTGTGATCTCTGCGCCAGCACAGATGCTTACCCCTGGAT 1560
Qy 1561 CCCAAGCTCACCGGTGTCTTCTGGGACCTTACCCCTTGGCATCGATCTTATTTGGAGC 1620
Db 1561 CCCAAGCTCACCGGTGTCTTCTGGGACCTTACCCCTTGGCATCGATCTTATTTGGAGC 1620
Qy 1621 CTGGCTGCCAACCACTTGAGTTCCTCAATCTCTTCAAGATGAAGATGTCGCTCATCTG 1680
Db 1621 CTGGCTGCCAACCACTTGAGTTCCTCAATCTCTTCAAGATGAAGATGTCGCTCATCTG 1680
Qy 1681 GGGTCTGTCACATGGCTTTGGGGTGTCTCTGAGTCTTCAACAGTGCCTTTGGC 1740
Db 1681 GGGTCTGTCACATGGCTTTGGGGTGTCTCTGAGTCTTCAACAGTGCCTTTGGC 1740
Qy 1741 CAGAGGCACCGGCTGCTGCTGGAGACGCTGCGGAGCTCACCTTCTGCTGGAGCTTTC 1800
Db 1741 CAGAGGCACCGGCTGCTGCTGGAGACGCTGCGGAGCTCACCTTCTGCTGGAGCTTTC 1800
Qy 1801 GGTACCTCTGTGTTCTTAGTCACTTCAAGTGGTGTGTGTGGGTGCGCAGGGCCGCC 1860
Db 1801 GGTACCTCTGTGTTCTTAGTCACTTCAAGTGGTGTGTGTGGGTGCGCAGGGCCGCC 1860
Qy 1861 TCGCCAGCATCTCATCTTCAATCAATGTTCTTCTTCCACAGCCCCAGCAAC 1920
Db 1861 TCGCCAGCATCTCATCTTCAATCAATGTTCTTCTTCCACAGCCCCAGCAAC 1920
Qy 1921 AGGCTGCTTACCCCCCGGAGGAGGTGCTCCAGGCCACGCTGGTGTGCTTGGCTTGGCC 1980
Db 1921 AGGCTGCTTACCCCCCGGAGGAGGTGCTCCAGGCCACGCTGGTGTGCTTGGCTTGGCC 1980

Qy 1981 ATGTGTCCTTCTGCTGCTTGGCACACCCCTGCACTGTGTGCAACCGCCACCGCGCGC 2040
Db 1981 ATGTGTCCTTCTGCTGCTTGGCACACCCCTGCACTGTGTGCAACCGCCACCGCGCGC 2040
Qy 2041 CTGCGGAGGAGGCCCTGACCGACAGAGGAGAAACAAGGCCGGGTTGCTGAGACTGCT 2100
Db 2041 CTGCGGAGGAGGCCCTGACCGACAGAGGAGAAACAAGGCCGGGTTGCTGAGACTGCT 2100
Qy 2101 GAGCATCTGTGAAATGCTGAGCTCCGATGAGGAAAGGAGGAGGCTTGGATGATGAA 2160
Db 2101 GAGCATCTGTGAAATGCTGAGCTCCGATGAGGAAAGGAGGAGGCTTGGATGATGAA 2160
Qy 2161 GAGGAGCCGAGCTGCTCCCTCCGAGTGTCTATGACACAGGCGCATCCACCATCGAG 2220
Db 2161 GAGGAGCCGAGCTGCTCCCTCCGAGTGTCTATGACACAGGCGCATCCACCATCGAG 2220
Qy 2221 TTCTGCTGGGCTGCGTCTCCACACCGCTTCTTACCTGCGCCCTGCGGCGCTGAGCCTG 2280
Db 2221 TTCTGCTGGGCTGCGTCTCCACACCGCTTCTTACCTGCGCCCTGCGGCGCTGAGCCTG 2280
Qy 2281 GCCACCGCCAGCTGCTCCGAGGTTCTGTGGGCCATGTGTGATGCGCATAGGCTTGGGCTG 2340
Db 2281 GCCACCGCCAGCTGCTCCGAGGTTCTGTGGGCCATGTGTGATGCGCATAGGCTTGGGCTG 2340
Qy 2341 GCGCGGAGGTGGCGGTGGCGGCTGTGGTGTGCTGCTCCCATCTTTGCGGCTTTGCGGCTG 2400
Db 2341 GCGCGGAGGTGGCGGTGGCGGCTGTGGTGTGCTGCTCCCATCTTTGCGGCTTTGCGGCTG 2400
Qy 2401 ATGACCGTGGCTATCTGCTGTGTGATGAGGAGGCTCTCAGCTTCTGCGAGCCCTGGG 2460
Db 2401 ATGACCGTGGCTATCTGCTGTGTGATGAGGAGGCTCTCAGCTTCTGCGAGCCCTGGG 2460
Qy 2461 CTGCACCTGGGTGGAATTCAGAAACAAGTTCTTACTCAGGCGAGGGCTTCAAGCTGAGTCCC 2520
Db 2461 CTGCACCTGGGTGGAATTCAGAAACAAGTTCTTACTCAGGCGAGGGCTTCAAGCTGAGTCCC 2520
Qy 2521 TTCACTTCTGCTGCGACAGATGACTAGGCGCCACTGCGAGTCTTCCGAGACTCTCTTCT 2580
Db 2521 TTCACTTCTGCTGCGACAGATGACTAGGCGCCACTGCGAGTCTTCCGAGACTCTCTTCT 2580
Qy 2581 GACCTCTGAGCGAGGAGGAGAAATAAGACGCTCCGCTCCGCTGCGCAAAAAA 2640
Db 2581 GACCTCTGAGCGAGGAGGAGAAATAAGACGCTCCGCTCCGCTGCGCAAAAAA 2640

RESULT 8

ABN96866

ID ABN96866 standard; DNA; 2655 BP.

AC ABN96866;

XX 13-AUG-2002 (first entry)

XX Gene #3364 used to diagnose liver cancer.

XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;

XX metastatic liver tumour; cytostatic; expression profile; disease state;

XX disease progression; drug toxicity; drug efficacy; drug metabolism.

XX Homo sapiens.

XX WO200229103-A2.

XX 11-APR-2002.

XX 02-OCT-2001; 2001WO-US030589.

XX 02-OCT-2000; 2000US-0237054P.

XX (GENE-) GENE LOGIC INC.

XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;

XX

WPI; 2002-426119/45.

Diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumor in a patient, involves detecting the level of expression of two or more genes in a liver tissue sample.

Claim 1; SEQ ID NO 3364; 298pp; English.

The invention relates to a novel method for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumour in a patient, and differentiating metastatic liver cancer from hepatocellular carcinoma in a patient, involving detecting the level of expression of two or more genes represented in ABN93503-ABN97455 in a tissue sample. The method of the invention has hepatotropic, and cytostatic activity. The method is useful for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma and metastatic liver carcinoma in a patient. The method is useful for identifying expression profiles which serve as useful diagnostic markers as well as markers that can be used to monitor disease states, disease progression, drug toxicity, drug efficacy and drug metabolism. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published/pct_sequences](http://wipo.int/pub/published/pct_sequences)

Sequence 2655 BP; 460 A; 885 C; 824 G; 486 T; 0 U; 0 Other;

Query Match 100.0%; Score 2640; DB 6; Length 2655;

Query Machine
Best Local Similarity

Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CGGCGTGC GGCAGCGGCAGCCAGCAGCGGAGCGCGGCAGCACACCCGGGACCATG 60

1 CGGCGTGC GGCACGGCAGCCAGCAGCGGCGGCGCAGCACACCCGGGACCATG 60

61 GGCTCCATGTTCCGAGCGAGAGGTGGCCCTGGTCCAGCTCTTTCTGCCCAACAGCGGCT 120

61 GGCTCCATGTTCCGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCCA CAGCGGCT 120

121 GCCTACACCTGCCGTGAGTCGGCTGGCGGAGCTGGGCCCTCGTGGAGTTCAGAGACCTCAAC 180

121 GCCTACACCTTGGGTGAGTCGGCTGGGGCAGCTTGGGCTCGTGGA GTTCAGAGACCTCAAC 180

181 GCCTCGGTGAGCGGCGCTTCCAGAGAGACGCTTTGTCGGTTGATGTTTGGCGCTGTGACGAGGCTG 240

[illegible][illegible]

243

2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

[illegible]

421 CTGCACAGCTGTAGCTTCCATCGTCCGCTGCTACGCTAGGGCCATGAACTTCAGCTGGCA 480

421 CTGCACCGCTGCAGCTCCACGCCGCCGCTGCTACGCCAGGGCCA TGAACCTCAGCTGGCA 480

481 GCCGCCACACAGATGGGGCTCAGAGAGGAGCCCTGCTCCAGGCCCCCGGGGGCCG 540

481 GCGGCCACACAGATGGGGCTTCAGAGAGGACGCCCTCTCCAGGCCCCCGGGGGCCG 540

541 CACCAGGACCTGAGGGTCAACTTTGTGGCAGGTGCCCTGGAGCCCCACAGGCCCTGCC 600

541 CACCAGGACCTGAGGGTCAACTTTGTGGCAGGTGCCCTGGAGCCCCACAAGGCCCTGCC 600

601 CTAGAGCGCCTGCTCTGGAGGGCCTGCCGCGGCTTCCTCATTTGCCAGCTTCAGGGAGCTG 660

301 CCAAGGGAGGCTGCCGGACCCCAACCCCGGACCTGCTGCATCCAGAGGAGACG 360
301 CCAAGGGAGGCTGCCGGACCCCAACCCCGGACCTGCTGCATCCAGAGGAGACG 360
361 GAGCGCTGGCCAGAGCTGGGGATGTGCGGGCAACAGAGGCCCTTGCGGCCAG 420
361 GAGCGCTGGCCAGAGCTGGGGATGTGCGGGCAACAGAGGCCCTTGCGGCCAG 420
421 CTGCAACAGCTCAGCTCCAGCCGCGCTGCTACGCGAGGCGCATGAACTCAGCTGGCA 480
421 CTGCAACAGCTCAGCTCCAGCCGCGCTGCTACGCGAGGCGCATGAACTCAGCTGGCA 480
481 GCGGCCCAACAGATGGGGCTCAGAGAGAGCCCTGCTCCAGGCCCGGGGGCG 540
481 GCGGCCCAACAGATGGGGCTCAGAGAGAGCCCTGCTCCAGGCCCGGGGGCG 540
541 CACAGGACCTGAGGCTCAACTTTGTGGCAGGTGCGGTGGAGCCCAAGGCCCTGCG 600
541 CACAGGACCTGAGGCTCAACTTTGTGGCAGGTGCGGTGGAGCCCAAGGCCCTGCG 600
601 CTAGAGCGCTGCTCTGAGAGGCGCTGCGCGGCTTCTCATTTGCCAGCTTCAGGGAGCTG 660
601 CTAGAGCGCTGCTCTGAGAGGCGCTGCGCGGCTTCTCATTTGCCAGCTTCAGGGAGCTG 660
661 GAGCAGCGCTGAGCACCCTGACGCGCGAGCCAGCCACGTGGATGACCTTCTCATC 720
661 GAGCAGCGCTGAGCACCCTGACGCGCGAGCCAGCCACGTGGATGACCTTCTCATC 720
721 TCTACTGGGCTGAGCAGATCGGACAGAAAGATCCGCAAGATCAAGGACTGCTTCACTGC 780
721 TCTACTGGGCTGAGCAGATCGGACAGAAAGATCCGCAAGATCAAGGACTGCTTCACTGC 780
781 CAGCTCTTCCGCTTCTGAGCAGAGGAGGCGCGCTCGGGCCCTGAGCAGCTGCA 840
781 CAGCTCTTCCGCTTCTGAGCAGAGGAGGCGCGCTCGGGCCCTGAGCAGCTGCA 840
841 CAGCAGAGCAGAGCTGAGAGGCTCTCGGGGAGACAGAGCGGCTTCTGAGCAGGCTG 900
841 CAGCAGAGCAGAGCTGAGAGGCTCTCGGGGAGACAGAGCGGCTTCTGAGCAGGCTG 900
901 CTAGGCGGGCTGCTGAGCTGCTGCGCCAGGGCAGGTGCAAGTCCAAAGATGAAGGCC 960
901 CTAGGCGGGCTGCTGAGCTGCTGCGCCAGGGCAGGTGCAAGTCCAAAGATGAAGGCC 960
961 GTGTACTGGCCCTGAAACAGTGACGCTGAGCACCAGCAAGTGCCTCATTCGCGAG 1020
961 GTGTACTGGCCCTGAAACAGTGACGCTGAGCACCAGCAAGTGCCTCATTCGCGAG 1020
1021 GCCTGGTCTGTGCGAGACTGCGCGCTGCGAGAGGCCCTGCGGACAGCTCGATG 1080
1021 GCCTGGTCTGTGCGAGACTGCGCGCTGCGAGAGGCCCTGCGGACAGCTCGATG 1080
1081 GAGGAGGAGTGAAGTGCCTGCTCAGCGCATCCCTGCGGGACATGCCCCCACTC 1140
1081 GAGGAGGAGTGAAGTGCCTGCTCAGCGCATCCCTGCGGGACATGCCCCCACTC 1140
1141 ATCCGCAACACCGCTTCAAGGCGAGCTTCAAGGCGATCGTGATGCTGAGGCTGGGC 1200
1141 ATCCGCAACACCGCTTCAAGGCGAGCTTCAAGGCGATCGTGATGCTGAGGCTGGGC 1200
1201 CGCTACAGGAGGTCAACCCCGCTCCCTACACCATCATCCTTCCCTTCTGTTTGT 1260
1201 CGCTACAGGAGGTCAACCCCGCTCCCTACACCATCATCCTTCCCTTCTGTTTGT 1260
1261 GTGATGTTCCGGGATGTGGCCACCGGCTGCTCATGTTCTTTCGCTTGGCCATGCTC 1320
1261 GTGATGTTCCGGGATGTGGCCACCGGCTGCTCATGTTCTTTCGCTTGGCCATGCTC 1320
1321 CTTGGGAGAACCGGCTGTGAAGCGCGGAGACGAGATCTGGCAGACTTCTTTC 1380
1321 CTTGGGAGAACCGGCTGTGAAGCGCGGAGACGAGATCTGGCAGACTTCTTTC 1380
1381 AGGGGCGCTACCTGCTCTGCTTATAGGCGCTGTTCTCCATCTACACCGGCTTCTACTAC 1440

1381 AGGGGCGCTACCTGCTCTGCTTATAGGCGCTGTTCTCCATCTACACCGGCTTCTACTAC 1440
1441 AACAGTGTCTCAGTTCGGGCCACAGCATTTTCCCTCGGGCTGGAGTGTGGCCGCAATG 1500
1441 AACAGTGTCTCAGTTCGGGCCACAGCATTTTCCCTCGGGCTGGAGTGTGGCCGCAATG 1500
1501 GCCAACCAAGTCTGGCTGGAGTATGATTTCTTGGGCCAGACACAGATGCTTACCCCTGGAT 1560
1501 GCCAACCAAGTCTGGCTGGAGTATGATTTCTTGGGCCAGACACAGATGCTTACCCCTGGAT 1560
1561 CCAACAGCTCACCGGTGTCTTCTGGGACCCCTACCCCTTTGGCATCGATCTTATTTGGAGC 1620
1561 CCAACAGCTCACCGGTGTCTTCTGGGACCCCTACCCCTTTGGCATCGATCTTATTTGGAGC 1620
1621 CTGGCTGCAACCACTTGAAGTCTCAACTCTCTTCAAGATGAAGATGTCGCTCATCTCTG 1680
1621 CTGGCTGCAACCACTTGAAGTCTCAACTCTCTTCAAGATGAAGATGTCGCTCATCTCTG 1680
1681 GCGCTGTCACATGGCCCTTTGGGGTGGTCTTGGAGTCTTCAACACAGTGCATTTGGC 1740
1681 GCGCTGTCACATGGCCCTTTGGGGTGGTCTTGGAGTCTTCAACACAGTGCATTTGGC 1740
1741 CAGAGGCACCGGCTGCTGTCGAGACGCTGCGGAGCTCACCTTCTGCTGGGACTCTTTC 1800
1741 CAGAGGCACCGGCTGCTGTCGAGACGCTGCGGAGCTCACCTTCTGCTGGGACTCTTTC 1800
1801 GGTACCTGCTGTTCTAGTCACTCAAGTGGTGTGTGTCGGCTGCCAGGGCGGC 1860
1801 GGTACCTGCTGTTCTAGTCACTCAAGTGGTGTGTGTCGGCTGCCAGGGCGGC 1860
1861 TCGCCAGCATCTCATCTTCAATCAAGTGTCTTCTTCTCCACAGCCCAAGCAAC 1920
1861 TCGCCAGCATCTCATCTTCAATCAAGTGTCTTCTTCTCCACAGCCCAAGCAAC 1920
1921 AGGCTGCTCTACCCCGGACAGAGTGGTTCAGGGCAAGCTGGTGGTCTGCTGGCTTGGCC 1980
1921 AGGCTGCTCTACCCCGGACAGAGTGGTTCAGGGCAAGCTGGTGGTCTGCTGGCTTGGCC 1980
1981 ATGCTGCCATCTGCTGCTTGGACACCCCTGCACTGCTGCAACCGCCGCGCCG 2040
1981 ATGCTGCCATCTGCTGCTTGGACACCCCTGCACTGCTGCAACCGCCGCGCCG 2040
2041 CTGCGGAGGCGCCGCTGACCGACAGGAGGAAACAGGGCCGGGTGCTGACCTTGCCT 2100
2041 CTGCGGAGGCGCCGCTGACCGACAGGAGGAAACAGGGCCGGGTGCTGACCTTGCCT 2100
2101 GACGATCTGTAATGGCTGGAGCTCCGATGAGGAAAGGCAAGGGGCTTGGATGATGAA 2160
2101 GACGATCTGTAATGGCTGGAGCTCCGATGAGGAAAGGCAAGGGGCTTGGATGATGAA 2160
2161 GAGGAGGCGAGCTGCTCCCTCCGAGTGTGTCACAGGCGCATCCACACCATCGAG 2220
2161 GAGGAGGCGAGCTGCTCCCTCCGAGTGTGTCACAGGCGCATCCACACCATCGAG 2220
2221 TTTCTGCTGGGCTGCTCTCAACACCGCTCTTACCTGCGCTGTGGGCCCTTGAGCCTG 2280
2221 TTTCTGCTGGGCTGCTCTCAACACCGCTCTTACCTGCGCTGTGGGCCCTTGAGCCTG 2280
2281 GCGCCAGCCAGCTGCTGCGAGTCTGTTGGGCGATGGTGAATGCGCATAGGCTTGGCCCTG 2340
2281 GCGCCAGCCAGCTGCTGCGAGTCTGTTGGGCGATGGTGAATGCGCATAGGCTTGGCCCTG 2340
2341 GCGCGGAGGTGGGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
2341 GCGCGGAGGTGGGCGTGGGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
2401 ATGACCGTGGCTATCTGCTGGTGAATGAGGAGCTCTCAGCCTTCTGACGCCCTTGGCG 2460
2401 ATGACCGTGGCTATCTGCTGGTGAATGAGGAGCTCTCAGCCTTCTGACGCCCTTGGCG 2460
2461 CTGCACTGGGTGAATTCAGAAAGTTCATCTCAGGCAAGGCTACAGCTGAGTCCC 2520

XX WPI; 2004-090520/09.
XX New composition comprising polynucleotide probes, useful as array
PT elements in a microarray for monitoring the expression of target
PT polynucleotides or purifying a subpopulation of mRNAs, cDNA, or genomic
PT fragments.
XX
PS Claim 6; SEQ ID NO 1094; 73pp; English.
XX
CC The invention relates to a composition of polynucleotide probes
CC comprising first polynucleotide probes comprising at least a portion of a
CC gene encoding a receptor-like polypeptide, second polynucleotide probes
CC comprising at least a portion of a gene encoding a transducing
CC polypeptide and third polynucleotide probes comprising at least a portion
CC of a gene encoding an effector-like polypeptide. The probes of the
CC composition are useful as array elements in a microarray for monitoring
CC the expression of target polynucleotides. The microarray is useful in the
CC diagnosis and treatment of cancer, an immunopathology or a
CC neuropathology. It can also be used for drug discovery and development,
CC toxicological and carcinogenicity studies, forensics or pharmacogenomics.
CC Microarrays can also be used for monitoring the progression of diseases
CC that may be associated with the altered expression of signalling pathway
CC polypeptides. The composition can also be used to purify a subpopulation
CC of mRNAs, cDNAs, or genomic fragments in a sample. The expression profile
CC is also useful for the diagnosis and treatment of cancer, e.g. cancers of
CC the adrenal gland, bladder, bone, bone marrow, brain, breast or cervix,
CC an immunopathology, e.g. AIDS, diabetes, pancreatitis, osteoporosis or
CC ulcerative colitis, or a neuropathology, e.g. dementia, amnesia,
CC epilepsy, Alzheimer's disease or depression. This sequence represents a
CC human polynucleotide probe of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 2655 BP; 460 A; 885 C; 824 G; 486 T; 0 U; 0 Other;

Query Match 100.0%; Score 2640; DB 12; Length 2655;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGCGTTCGGGACGGGACGACGAGCGGGCGGCGGACACACACCGGGACCATG 60
DB |||||
DB 1 CGGCGTTCGGGACGGGACGACGAGCGGGCGGCGGACACACCGGGACCATG 60
QY 61 GGCTCCATGTTCCGAGCGAGGAGGTGGCCCTCGTCCAGCTCTTTCTGCCACACAGCGGCT 120
DB |||||
DB 61 GGCTCCATGTTCCGAGCGAGGAGGTGGCCCTCGTCCAGCTCTTTCTGCCACACAGCGGCT 120
QY 121 GCCTACACCTGCGTGGTGGGCGAGCTGGGCGCTCGTGAGTTTCAGAGACCTCAAC 180
DB |||||
DB 121 GCCTACACCTGCGTGGTGGGCGAGCTGGGCGCTCGTGAGTTTCAGAGACCTCAAC 180
QY 181 GCCTCGGTGAGCGCTTTCAGAGACGCTTTGTGGTTCATGTTTGGCGCTGTGAGGAGCTG 240
DB |||||
DB 181 GCCTCGGTGAGCGCTTTCAGAGACGCTTTGTGGTTCATGTTTGGCGCTGTGAGGAGCTG 240
QY 241 GAGAAGACCTTCACTTCTCCAGGAGGAGGTGGCGGCTGGGCTGGTCTGCCCGCG 300
DB |||||
DB 241 GAGAAGACCTTCACTTCTCCAGGAGGAGGTGGCGGCTGGGCTGGTCTGCCCGCG 300
QY 301 CCAAGGGGAGGCTGCGGACACCCCGGACCTGCTGGCGATCCAGAGGAGACG 360
DB |||||
DB 301 CCAAGGGGAGGCTGCGGACACCCCGGACCTGCTGGCGATCCAGAGGAGACG 360
QY 361 GAGCGCTGCGGACGAGCTGGGAGTGTGGGCGCAACACGAGGCGCTGCGGGCCAG 420
DB |||||
DB 361 GAGCGCTGCGGACGAGCTGGGAGTGTGGGCGCAACACGAGGCGCTGCGGGCCAG 420
QY 421 CTCACACAGCTGCAGCTTCAACCGCGCGTGTCTACGCCAGGCGCATGAACCTCAGCTGGCA 480
DB |||||
DB 421 CTCACACAGCTGCAGCTTCAACCGCGCGTGTCTACGCCAGGCGCATGAACCTCAGCTGGCA 480

QY 481 GCCGCCACACACAGATGGGGCTCAGAGAGGAGCGCCCTGCTCCAGAGCCCCCGGGGGGGCGG 540
DB |||||
DB 481 GCCGCCACACACAGATGGGGCTCAGAGAGGAGCGCCCTGCTCCAGAGCCCCCGGGGGGGCGG 540
QY 541 CACCAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGCGTGGAGCGCCCAACAGGCGCCCTGCC 600
DB |||||
DB 541 CACCAGGACCTGAGGGTCAACTTTGTGGCAGGTGCGCGTGGAGCGCCCAACAGGCGCCCTGCC 600
QY 601 CTAGAGCGCTGCTCTCGAGGGCTGCGCGGGCTTCTCTCATTTGCGAGCTTCAGGAGCTG 660
DB |||||
DB 601 CTAGAGCGCTGCTCTCGAGGGCTGCGCGGGCTTCTCTCATTTGCGAGCTTCAGGAGCTG 660
QY 661 GAGCAGCGCTGGAGCACCCCGTGCACGGGCGAGCAGCAGCTGGATGACCTTCTCATC 720
DB |||||
DB 661 GAGCAGCGCTGGAGCACCCCGTGCACGGGCGAGCAGCAGCTGGATGACCTTCTCATC 720
QY 721 TCCTACTGGGCTGAGCAGATCGGACAGAATCGCAAGATCACGGACTGCTTCCACTGC 780
DB |||||
DB 721 TCCTACTGGGCTGAGCAGATCGGACAGAATCGCAAGATCACGGACTGCTTCCACTGC 780
QY 781 CACGTCTTCCGTTTCTGACAGGAGGAGGCGCCCTCGGGGCGCTTCGACAGCTGCA 840
DB |||||
DB 781 CACGTCTTCCGTTTCTGACAGGAGGAGGCGCCCTCGGGGCGCTTCGACAGCTGCA 840
QY 841 CAGCAGAGCAGGAGCTGACAGGAGTCTCTCGGGGAGACAGAGCGGTTCTCTGAGCCAGGTG 900
DB |||||
DB 841 CAGCAGAGCAGGAGCTGACAGGAGTCTCTCGGGGAGACAGAGCGGTTCTCTGAGCCAGGTG 900
QY 901 CTAGGCGGCTGCTGACAGTGTGCGCCAGGCGAGGTGCGAGTCCACAGATCAAGGCC 960
DB |||||
DB 901 CTAGGCGGCTGCTGACAGTGTGCGCCAGGCGAGGTGCGAGTCCACAGATCAAGGCC 960
QY 961 GTGTACTGCGCTGAAACCAAGTGCAGCGTGCAGCACCAAGTGCCTCATTTGCCGAG 1020
DB |||||
DB 961 GTGTACTGCGCTGAAACCAAGTGCAGCGTGCAGCACCAAGTGCCTCATTTGCCGAG 1020
QY 1021 GCCTGTGCTCTGTCGGAGACCTGCGCGCGCTGACAGAGGCGCTGCGGACAGCTCATG 1080
DB |||||
DB 1021 GCCTGTGCTCTGTCGGAGACCTGCGCGCGCTGACAGAGGCGCTGCGGAGCAGCTCATG 1080
QY 1081 GAGGAGGAGTGTGCTGCGGTGCTCACCGCATCCCTGCGGGACATGCCCCCACAATC 1140
DB |||||
DB 1081 GAGGAGGAGTGTGCTGCGGTGCTCACCGCATCCCTGCGGGACATGCCCCCACAATC 1140
QY 1141 ATCCGCAACCAACCGCTTTCAGCGGACGTTCCAGGGACATCGTGGATCGCTACGGCGTGGGC 1200
DB |||||
DB 1141 ATCCGCAACCAACCGCTTTCAGCGGACGTTCCAGGGACATCGTGGATCGCTACGGCGTGGGC 1200
QY 1201 CGCTACAGAGGTCAACCCCGCTCCCTACACATCATCACTTCCCTCTCTCTGTTGCT 1260
DB |||||
DB 1201 CGCTACAGAGGTCAACCCCGCTCCCTACACATCATCACTTCCCTCTCTCTGTTGCT 1260
QY 1261 GTGATGTTGGGAGTGTGGGCGACGGGCTGCTCATGTTCTCTTGGCCCTTGGCCCATGTC 1320
DB |||||
DB 1261 GTGATGTTGGGAGTGTGGGCGACGGGCTGCTCATGTTCTCTTGGCCCTTGGCCCATGTC 1320
QY 1321 CTTTGGGAGAACCGAGCGGCTGTGAAGCGCGCAGAACAGAGATCTGGCAGACTTCTTC 1380
DB |||||
DB 1321 CTTTGGGAGAACCGAGCGGCTGTGAAGCGCGCAGAACAGAGATCTGGCAGACTTCTTC 1380
QY 1381 AGGGGCGCTTACCTGCTCTGCTTATGGGCTGTTCTTCATCTACACGGGCTTCATCTAC 1440
DB |||||
DB 1381 AGGGGCGCTTACCTGCTCTGCTTATGGGCTGTTCTTCATCTACACGGGCTTCATCTAC 1440
QY 1441 AACGAGTGTTCAGTTCGCGCACACAGCATCTTCCCTCGGGCTGGAGTGGCGCCCATG 1500
DB |||||
DB 1441 AACGAGTGTTCAGTTCGCGCACACAGCATCTTCCCTCGGGCTGGAGTGGCGCCCATG 1500
QY 1501 GCCAACAGCTGTGGCTGGATGATTCCTTGGCCAGCACAGATGCTTACCCCTGGAT 1560
DB |||||
DB 1501 GCCAACAGCTGTGGCTGGATGATTCCTTGGCCAGCACAGATGCTTACCCCTGGAT 1560
QY 1561 CCCAACGTACCGGCTGTCTTCTTCTGGGACCCCTACCCCTTTGGGATCGATTCCTATTTGGAGC 1620

| | | | | | | | | | | | | |
|----|------|-------|---------|-----------|------------|--------|-----------|----------|---------|----------|---------|------|
| Db | 1561 | | CCCAA | CGTCA | CCGGTGTCTT | CTCTGG | GAACCTT | ACCCCTT | GGCATCG | ATTTGG | GAGC | 1620 |
| Qy | 1621 | CTGG | CTGCA | CAACCACTT | CTGAGCTT | CTCTAA | CTCTTCA | AGATGA | AGATGT | CGCTCA | CTCTG | 1680 |
| Db | 1621 | CTGG | CTGCA | CAACCACTT | GAGCTT | CTCTAA | CTCTTCA | AGATGA | AGATGT | CGCTCA | CTCTG | 1680 |
| Qy | 1681 | GGCG | CTGCA | CATGGCCTT | GGGGTGT | CTCTCG | AGACTT | CTTCA | ACACAG | CTGCACTT | GGC | 1740 |
| Db | 1681 | GGCG | CTGCA | CATGGCCTT | GGGGTGT | CTCTCG | AGACTT | CTTCA | ACACAG | CTGCACTT | GGC | 1740 |
| Qy | 1741 | CAGAG | GCACCG | CTGCTCTG | GAGACCG | CTGCGG | AGCTCA | CTTCTCTG | CTGCGAG | CTTCTT | C | 1800 |
| Db | 1741 | CAGAG | GCACCG | CTGCTCTG | GAGACCG | CTGCGG | AGCTCA | CTTCTCTG | CTGCGAG | CTTCTT | C | 1800 |
| Qy | 1801 | GTTT | ACCTCG | TGTTCTT | AGTCATCT | TACA | AGTGGCTGT | GTCTCG | GGCTGCC | AGGCGCC | C | 1860 |
| Db | 1801 | GTTT | ACCTCG | TGTTCTT | AGTCATCT | TACA | AGTGGCTGT | GTCTCG | GGCTGCC | AGGCGCC | C | 1860 |
| Qy | 1861 | TCGCC | AGCATCT | CTCAT | CCACTT | CTCA | ACAATGTT | CTCTTCT | CTCCAC | AGCCCC | CAGCAAC | 1920 |
| Db | 1861 | TCGCC | AGCATCT | CTCAT | CCACTT | CTCA | ACAATGTT | CTCTTCT | CTCCAC | AGCCCC | CAGCAAC | 1920 |
| Qy | 1921 | AGGCT | GTCTA | CCCCCGG | CAGGAGT | TGTC | CAGGCC | ACGCTG | TGTCTT | CGCTG | GGCTTGG | 1980 |
| Db | 1921 | AGGCT | GTCTA | CCCCCGG | CAGGAGT | TGTC | CAGGCC | ACGCTG | TGTCTT | CGCTG | GGCTTGG | 1980 |
| Qy | 1981 | ATG | GTGCC | ATCTCT | GTCTTGG | CACAC | CCCTG | CACTG | CTGCA | CGCC | ACCGCG | 2040 |
| Db | 1981 | ATG | GTGCC | ATCTCT | GTCTTGG | CACAC | CCCTG | CACTG | CTGCA | CGCC | ACCGCG | 2040 |
| Qy | 2041 | CTCG | AGGAG | CGCCCG | CTGAC | CCGAC | GAGG | AAAAA | AGGCGGG | TGCTGG | ACCTTGC | 2100 |
| Db | 2041 | CTCG | AGGAG | CGCCCG | CTGAC | CCGAC | GAGG | AAAAA | AGGCGGG | TGCTGG | ACCTTGC | 2100 |
| Qy | 2101 | GAC | GCATCT | GTGA | TGGCTG | AGCTCC | GA | TCGAT | GAG | AAAAA | AGGCGGG | 2160 |
| Db | 2101 | GAC | GCATCT | GTGA | TGGCTG | AGCTCC | GA | TCGAT | GAG | AAAAA | AGGCGGG | 2160 |
| Qy | 2161 | GAG | GAG | CGGAG | CTCGT | CCCCCT | CGAG | TGCTCAT | GCA | CAGG | CCATCC | 2220 |
| Db | 2161 | GAG | GAG | CGGAG | CTCGT | CCCCCT | CGAG | TGCTCAT | GCA | CAGG | CCATCC | 2220 |
| Qy | 2221 | TTT | CTGG | CTGG | CTCGT | CTCA | ACCG | CTCTT | ACCTG | CGCCTG | TGGG | 2280 |
| Db | 2221 | TTT | CTGG | CTGG | CTCGT | CTCA | ACCG | CTCTT | ACCTG | CGCCTG | TGGG | 2280 |
| Qy | 2281 | GCC | ACCG | CCAG | CTGT | CGGAG | TTCT | GTGG | CCATG | TGTG | ATGCG | 2340 |
| Db | 2281 | GCC | ACCG | CCAG | CTGT | CGGAG | TTCT | GTGG | CCATG | TGTG | ATGCG | 2340 |
| Qy | 2341 | GGC | CGGAG | TGGG | CTGG | CGG | CTGT | GTGT | TCC | CTTCTT | TGCG | 2400 |
| Db | 2341 | GGC | CGGAG | TGGG | CTGG | CGG | CTGT | GTGT | TCC | CTTCTT | TGCG | 2400 |
| Qy | 2401 | ATG | ACCG | TGGCTAT | CTCTG | CTGGT | GATG | GA | AGGACT | CTC | CAG | 2460 |
| Db | 2401 | ATG | ACCG | TGGCTAT | CTCTG | CTGGT | GATG | GA | AGGACT | CTC | CAG | 2460 |
| Qy | 2461 | CTG | CAC | TGGG | TGGA | TTTCC | AGAAC | AA | AGTTCT | TACT | CAGG | 2520 |
| Db | 2461 | CTG | CAC | TGGG | TGGA | TTTCC | AGAAC | AA | AGTTCT | TACT | CAGG | 2520 |
| Qy | 2521 | TT | CAC | TTCTG | CTGCC | ACAG | ATGA | CTTAG | GGCC | CACTG | CAG | 2580 |
| Db | 2521 | TT | CAC | TTCTG | CTGCC | ACAG | ATGA | CTTAG | GGCC | CACTG | CAG | 2580 |
| Qy | 2581 | GAC | CTCTG | AG | CGAGG | AG | ATGA | AG | CGGT | TCG | CG | 2640 |
| Db | 2581 | GAC | CTCTG | AG | CGAGG | AG | ATGA | AG | CGGT | TCG | CG | 2640 |

Db 2401 ATGACCGTGCTATCTCTGCTGGTGATGAGGGGACTCTCAGCCCTTCTGCGACGCCCTGCGG 2460
Qy 2461 CTGCACTGGGTGAATTCAGAAACAAGTTCTACTCAGGCACGGGCTTACAAGCTGAGTCCC 2520
Db 2461 CTGCACTGGGTGAATTCAGAAACAAGTTCTACTCAGGCACGGGCTTACAAGCTGAGTCCC 2520
Qy 2521 TTCACTCTGCTGCCACAGATGACTAGGGCCCACTGCAGTCTCTGCGACACCTCTTCTT 2580
Db 2521 TTCACTCTGCTGCCACAGATGACTAGGGCCCACTGCAGTCTCTGCGACACCTCTTCTT 2580
Qy 2581 GACCTCTGAGGCGAGAGAGGAATAAAGACGGTCCGCCCTGGCAAAAAAAAAAAAAA 2640
Db 2581 GACCTCTGAGGCGAGAGAGGAATAAAGACGGTCCGCCCTGGCAAAAAAAAAAAAAA 2640

RESULT 13

AD128819
ID AD128819 standard; cDNA; 2700 BP.

XX AC AD128819;

XX DT 22-APR-2004 (first entry)

XX DE Human modifier of p53 (MP53) gene TCIRG-1.

XX KW Human; MP53; modifier of p53; p53; cytosolic; gene therapy;

XX KW T-cell immune regulator 1; TCIRG-1; gene; ss.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

XX FT CDS 109..2601

XX FT /tag= a

XX FT /product= "MP53 TCIRG-1"

XX WO2004004766-A1.

XX PD 15-JAN-2004.

XX PF 09-JUL-2003; 2003WO-US021378.

XX PR 10-JUL-2002; 2002US-0394992P.

XX PR 07-AUG-2002; 2002US-0401604P.

XX PR 16-SEP-2002; 2002US-0410988P.

XX PR 25-NOV-2002; 2002US-0428837P.

XX (EXEL-) EXELIXIS INC.

XX Costa MA, Maxwell ME, Lackner MR, Hung T, O'Brien CL, Jin Y;

XX Nicoll M, Hai B, Zhang H, Lickteig K, Amundsen CD;

XX WPI; 2004-142922/14.

XX P-PSDB; AD128834.

XX REFSEQ; NM_006019.2.

XX Identifying a candidate p53 pathway modulating agent for treating e.g., cancer, comprises contacting an assay system comprising a MP53 polypeptide or nucleic acid with a test agent and detecting a test agent-biased activity.

XX Example 2; SEQ ID NO 10; 139pp; English.

XX In the present invention, genetic screens were designed to identify modifiers of the p53 pathway in *Caenorhabditis elegans*, where a homozygous p53 deletion mutant was used. Various specific genes were silenced by RNA inhibition. Genes causing altered phenotypes in the worms were identified as modifiers of the p53 pathway. Human orthologs, denoted as modifiers of p52 (MP53), of these modifiers were identified. These include the present sequence, characterised as encoding T-cell immune regulator 1, ATPase, H⁺ transporting, or lysosomal V0 protein a isoform 3. MP53 nucleic acids and polypeptides are attractive drug targets for the treatment of pathologies associated with a defective p53 signalling pathway, such as cancer. Methods for modulating MP53 function and/or the

CC p53 pathway in a mammalian cell involve contacting the cell with an agent that specifically binds a MP53 polypeptide or nucleic acid. The agent may be a small molecular modulator, a nucleic acid modulator or an antibody.

XX Sequence 2700 BP; 451 A; 905 C; 850 G; 494 T; 0 U; 0 Other;

Query Match 81.1%; Score 2142; DB 12; Length 2700;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2582; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

Qy 38 GCGCAGCACACCGGGGACCATGGGCTCCATGTTCCGAGCGAGGAGGTGGCCCTGGTCC 97

Db 89 GCGCAGCACACCGGGGACCATGGGCTCCATGTTCCGAGCGAGGAGGTGGCCCTGGTCC 148

Qy 98 AGCTCTTTCTGCCCCACAGCGGCTGCTACACCTGCGTGAAGTGGCTGGGGGAGCTGGGCC 157

Db 149 AGCTCTTTCTGCCCCACAGCGGCTGCTACACCTGCGTGAAGTGGCTGGGGGAGCTGGGCC 208

Qy 158 TCGTGGAGTTTCAGAGACCTCAACGCTCGGTGAGCGCTTCCAGAGACGCTTTGTGTTG 217

Db 209 TCGTGGAGTTTCAGAGACCTCAACGCTCGGTGAGCGCTTCCAGAGACGCTTTGTGTTG 268

Qy 218 ATGTTTGGCGCTGTGAGGAGCTGGAGAGACCTTCACTTCTTCGAGGAGGAGTGGCGC 277

Db 269 ATGTTTGGCGCTGTGAGGAGCTGGAGAGACCTTCACTTCTTCGAGGAGGAGTGGCGC 328

Qy 278 GGGCTGGGCTGTGCTGCCCCGCCAAAGGGAGGCTGCGGACACCCACCCCGGGACC 337

Db 329 GGGCTGGGCTGTGCTGCCCCGCCAAAGGGAGGCTGCGGACACCCACCCCGGGACC 388

Qy 338 TCGTGGCATCCAGGAGGAGCGAGCGCTGCGCCAGGAGCTGCGGGAGTGTGCGGGCA 397

Db 389 TCGTGGCATCCAGGAGGAGCGAGCGCTGCGCCAGGAGCTGCGGGAGTGTGCGGGCA 448

Qy 398 ACCAGCAGCCCTGCGGGGCCAGCTGACAGCTGCACAGCTCCACGCCCGCTGTACTACGCC 457

Db 449 ACCAGCAGCCCTGCGGGGCCAGCTGACAGCTGCACAGCTCCACGCCCGCTGTACTACGCC 508

Qy 458 AGGGCCATGAACCTCAGCTGGCAGCGGCCCAACAGATGGGGCTTCAGAGAGACGCCCC 517

Db 509 AGGGCCATGAACCTCAGCTGGCAGCGGCCCAACAGATGGGGCTTCAGAGAGACGCCCC 568

Qy 518 TGTCTCAGGCCCCCGGGGGCGGCACACAGGACTGAGGGGTCAACTTTGTGGCAGGTGCCG 577

Db 569 TGTCTCAGGCCCCCGGGGGCGGCACACAGGACTGAGGGGTCAACTTTGTGGCAGGTGCCG 628

Qy 578 TGGAGCCCCACAAAGGCCCTGCTGTAGAGCGCTGCTTGGAGGGCTGCTGCGGGTTC 637

Db 629 TGGAGCCCCACAAAGGCCCTGCTGTAGAGCGCTGCTTGGAGGGCTGCTGCGGGTTC 688

Qy 638 TCATTTGCCAGCTTCAGGAGCTGGAGCAGCGCTGGAGCACCCCTGACGGGGAGCCAG 697

Db 689 TCATTTGCCAGCTTCAGGAGCTGGAGCAGCGCTGGAGCACCCCTGACGGGGAGCCAG 748

Qy 698 CCAGCTGAGTGAACCTTCT 757

Db 749 CCAGCTGAGTGAACCTTCT 808

Qy 758 AGATCAACGAGCTGCTTCCATCCACAGCTTCCCGTTTCTGACGACGAGGAGGCCGCC 817

Db 809 AGATCAACGAGCTGCTTCCATCCACAGCTTCCCGTTTCTGACGACGAGGAGGCCGCC 868

Qy 818 TCGGGGCCCTGACGACGCTGCAACAGCAGACGACGAGCTGACGAGGAGTCTCTCGGGAGA 877

Db 869 TCGGGGCCCTGACGACGCTGCAACAGCAGACGACGAGCTGACGAGGAGTCTCTCGGGAGA 928

Qy 878 CAGAGCGGTCTCTGACGAGCTGCTAGGCGGGTGTCTGAGCTGCTGCGCCACGAGGCGAGG 937

Db 929 CAGAGCGGTCTCTGACGAGCTGCTAGGCGGGTGTCTGAGCTGCTGCGCCACGAGGCGAGG 988

Qy 938 TGCAGTCCACAGATGAAGGCGGTGTACTGCGCCCTGAAACAGTGCAGCGTGAACCA 997

Db 989 TGCAGTCCACAGATGAAGGCGGTGTACTGCGCCCTGAAACAGTGCAGCGTGAACCA 1048

QY 998 CGCACAAGTGGCTCATTTGCCAGGCGCTGGTCTCTGTGCGAGACCTGCCCCCGCTGCGAG 1057
Db |||||
QY 1049 CGCACAAGTGGCTCATTTGCCAGGCGCTGGTCTCTGTGCGAGACCTGCCCCCGCTGCGAG 1108
Db |||||
QY 1058 AGGCGCTTGCAGGACAGCTTCGATGAGAGGAGGAGTGAAGTGGCTCACCAGCATCCCTCT 1117
Db |||||
QY 1109 AGGCGCTTGCAGGACAGCTTCGATGAGAGGAGGAGTGAAGTGGCTCACCAGCATCCCTCT 1168
Db |||||
QY 1118 GCCGGGACATGCCCGCCACACATCATTCGCAACAAACCGCTTACCGGCGAGCTTCCAGGGCA 1177
Db |||||
QY 1169 GCCGGGACATGCCCGCCACACATCATTCGCAACAAACCGCTTACCGGCGAGCTTCCAGGGCA 1228
Db |||||
QY 1178 TCGTGGATCGCTACGGCGTGGCGCGCTACACAGGAGGTCAACCCCGCTCCCTACACCATCA 1237
Db |||||
QY 1229 TCGTGGATCGCTACGGCGTGGCGCGCTACACAGGAGGTCAACCCCGCTCCCTACACCATCA 1288
Db |||||
QY 1238 TCACCTTCCCTTCTCTGTTCTGTGATGTTTGGGGATGTGGGCCACCGGGCTGCTCATGT 1297
Db |||||
QY 1289 TCACCTTCCCTTCTCTGTTCTGTGATGTTTGGGGATGTGGGCCACCGGGCTGCTCATGT 1348
Db |||||
QY 1298 TCCTCTTGGCGCTGGCGCATGTCCTTGGGGAACCGACCGGCTGTGAAAGCCGCGCAGA 1357
Db |||||
QY 1349 TCCTCTTGGCGCTGGCGCATGTCCTTGGGGAACCGACCGGCTGTGAAAGCCGCGCAGA 1408
Db |||||
QY 1358 ACGAGATCTGGCAGACTTCTTACGGGGCGCTACCTGCTCTGCTTATGGGCGCTGTCT 1417
Db |||||
QY 1409 ACGAGATCTGGCAGACTTCTTACGGGGCGCTACCTGCTCTGCTTATGGGCGCTGTCT 1468
Db |||||
QY 1418 CCATCTACACCGGCTTCTATCTACAAACGAGTCTTTCAGTTCGGCGCCACACAGCATCTTCCCT 1477
Db |||||
QY 1469 CCATCTACACCGGCTTCTATCTACAAACGAGTCTTTCAGTTCGGCGCCACACAGCATCTTCCCT 1528
Db |||||
QY 1478 CGGCTGGAGTGTGGCGCCCATGCCAAACAGTCTGGCTGGAGTGAATCTTCTGGGCC 1537
Db |||||
QY 1529 CGGCTGGAGTGTGGCGCCCATGCCAAACAGTCTGGCTGGAGTGAATCTTCTGGGCC 1588
Db |||||
QY 1538 AGCACAGATCTTACCTGGATCCCAACGTCACCGGTGCTTCTTGGGACCTTACCGCT 1597
Db |||||
QY 1589 AGCACAGATCTTACCTGGATCCCAACGTCACCGGTGCTTCTTGGGACCTTACCGCT 1648
Db |||||
QY 1598 TTGGCATCGATCTCTATTTGGAGCTGTGCTGCAACCACTTTCAGTCTTCTCAACTCTTCA 1657
Db |||||
QY 1649 TTGGCATCGATCTCTATTTGGAGCTGTGCTGCAACCACTTTCAGTCTTCTCAACTCTTCA 1708
Db |||||
QY 1658 AGATGAAGATGTCCGTTCATCTTGGCGCTGTGTCACATGGCTTTGGGTGTCTCTCGAG 1717
Db |||||
QY 1709 AGATGAAGATGTCCGTTCATCTTGGCGCTGTGTCACATGGCTTTGGGTGTCTCTCGAG 1768
Db |||||
QY 1718 TCTTCAACAGTGTGACCTTTGGCCAGAGGACACCGGCTGTGCTGGAGACGCTGCCGAGC 1777
Db |||||
QY 1769 TCTTCAACAGTGTGACCTTTGGCCAGAGGACACCGGCTGTGCTGGAGACGCTGCCGAGC 1828
Db |||||
QY 1778 TCACCTTCTCTGCTGGGACTCTTTCGGTTACCTCTGTTTCCCTAGTTCATCTACAAGTGGCTGT 1837
Db |||||
QY 1829 TCACCTTCTCTGCTGGGACTCTTTCGGTTACCTCTGTTTCCCTAGTTCATCTACAAGTGGCTGT 1888
Db |||||
QY 1838 GTGTCTGGGTGTCAGGCGCGCTCTGCG --- CCCAGCATCTCTCATCCACTTTCATCAACATGT 1894
Db |||||
QY 1889 GTGTCTGGGTGTCAGGCGCGCTCTGCGCGCCCGCCAGCATCTCTCATCAACATGT 1948
Db |||||
QY 1895 TCCTCTTCTCCACAGGCGCCAGCAACAGGCTGTCTACCCCGCGCAGAGGTGTCTCAGG 1954
Db |||||
QY 1949 TCCTCTTCTCCACAGGCGCCAGCAACAGGCTGTCTACCCCGCGCAGAGGTGTCTCAGG 2008
Db |||||
QY 1955 CCAGCTGTGTGCTCTGGCTTGGCCATGTGGCCATGTCCTGCTCTTGGCACACCCCTGC 2014
Db |||||
QY 2009 CCAGCTGTGTGCTCTGGCTTGGCCATGTGGCCATGTCCTGCTCTTGGCACACCCCTGC 2068
Db |||||
QY 2015 ACCTGTGTGACCGCCACCGCGCGCTTGGGAGGAGGCGCGCTGTGACCGCAGAGGAGAA 2074
Db |||||
QY 2069 ACCTGTGTGACCGCCACCGCGCGCTTGGGAGGAGGCGCGCTGTGACCGCAGAGGAGAA 2128
Db |||||

QY 2075 ACAAGGCGGGTTCTGAGACCTGCTGACGATCTGTGAATGGCTGGAGCTCCGATGAGG 2134
Db |||||
QY 2129 ACAAGGCGGGTTCTGAGACCTGCTGACGATCTGTGAATGGCTGGAGCTCCGATGAGG 2188
Db |||||
QY 2135 AAAAGGACGGGGCTGATGATGAAGAGAGGCGGAGCTGTCCCTCCGAGGTGCTCA 2194
Db |||||
QY 2189 AAAAGGACGGGGCTGATGATGAAGAGAGGCGGAGCTGTCCCTCCGAGGTGCTCA 2248
Db |||||
QY 2195 TGCACAGGCGCATCCACACCATCGAGTTCTGCTGGGCTGCTTCCAAACACCGCTCT 2254
Db |||||
QY 2249 TGCACAGGCGCATCCACACCATCGAGTTCTGCTGGGCTGCTTCCAAACACCGCTCT 2308
Db |||||
QY 2255 ACCTGCGCTGTGGGCGCTGAGCTGCGCCACGCGCCAGCTCTCGAGGTTCTGTGGGCA 2314
Db |||||
QY 2309 ACCTGCGCTGTGGGCGCTGAGCTGCGCCACGCGCCAGCTCTCGAGGTTCTGTGGGCA 2368
Db |||||
QY 2315 TGTGATGCGCATAGGCTGCGGCTGCGGCGGAGGTGGGCGTGGCGCTGTGGTGTGG 2374
Db |||||
QY 2369 TGTGATGCGCATAGGCTGCGGCTGCGGCGGAGGTGGGCGTGGCGCTGTGGTGTGG 2428
Db |||||
QY 2375 TCCCATCTTTGCGCGCTTTGCGGTGATGACCGTGGCTATCTCTGCTGTGATGAGGAC 2434
Db |||||
QY 2429 TCCCATCTTTGCGCGCTTTGCGGTGATGACCGTGGCTATCTCTGCTGTGATGAGGAC 2488
Db |||||
QY 2435 TCTCAGCTTCTGCGCGCTGCGGCTGCGCTGCGGCTGCGCTGCGGCTGCGGCTGCGGCT 2494
Db |||||
QY 2489 TCTCAGCTTCTGCGCGCTGCGGCTGCGCTGCGGCTGCGCTGCGGCTGCGGCTGCGGCT 2548
Db |||||
QY 2495 CAGGACGGGCTACAGCTGAGTCCCTTTCAGCTTTCGCTGCGCAGATGACTAGGCGCCAC 2554
Db |||||
QY 2549 CAGGACGGGCTACAGCTGAGTCCCTTTCAGCTTTCGCTGCGCAGATGACTAGGCGCCAC 2608
Db |||||
QY 2555 TGCAGGTCTGCGCAGACCTCTTCTGACCTCTGAGGCGAGAGGAAATAAGACGGTCC 2614
Db |||||
QY 2609 TGCAGGTCTGCGCAGACCTCTTCTGACCTCTGAGGCGAGAGGAAATAAGACGGTCC 2668
Db |||||
QY 2615 GCCCTGGCA 2623
Db |||||
QY 2669 GCCCTGGCA 2677
Db |||||

RESULT 14

ADP10401
ID ADP10401 standard; DNA; 2700 BP.

XX ADP10401;

XX AC
XX AC
XX 12-AUG-2004 (first entry)

XX DE Reference mRNA sequences for marker probe #78.

XX DE transplant rejection; immune system; rheumatoid arthritis; lupus;

XX KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS; ss.

XX OS Homo sapiens.

XX FN WO2004042346-A2.

XX PD 21-MAY-2004.

XX PF 24-APR-2003; 2003WO-US012946.

XX XX 24-APR-2002; 2002US-00131831.

XX PR 20-DEC-2002; 2002US-00325899.

XX XX (EXPR-) EXPRESSION DIAGNOSTICS INC.

XX XX Wohlgenuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;

XX PI Rosenberg S;

XX XX WPI; 2004-400724/37.

XX PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,

| | | | | | | |
|----|------|---|------|------|---|------|
| QY | 338 | TGCTGCGCATCCAGGAGGACGAGCGCCTGSCCAGGAGCTGCGGGATGTGCGGGGCA | 397 | 1418 | CCATCTACACCGCGCTTCATCTACAAAGAGTGTTCAGTCGCGCCACAGCATCTCCCCCT | 1477 |
| Db | 389 | TGCTGCGCATCCAGGAGGACGAGCGCCTGSCCAGGAGCTGCGGGATGTGCGGGGCA | 448 | 1469 | CCATCTACACCGCGCTTCATCTACAAAGAGTGTTCAGTCGCGCCACAGCATCTCCCCCT | 1528 |
| QY | 398 | ACCAGAGGCGCTTCGCGGGCCAGCTGCACAGCTGCAGCTCCACGCCCGCGTGTACGCC | 457 | 1478 | CGGGCTGGAGTGTGGCGGCATGGCCAAACAGTCTGGCTGGAGTGATGCAATTCCTGGCCCC | 1537 |
| Db | 449 | ACCAGAGGCGCTTCGCGGGCCAGCTGCACAGCTGCAGCTCCACGCCCGCGTGTACGCC | 508 | 1529 | CGGGCTGGAGTGTGGCGGCATGGCCAAACAGTCTGGCTGGAGTGATGCAATTCCTGGCCCC | 1588 |
| QY | 458 | AGGGCCATGAACTCAGCTGGGAGCGGCCACACAGATGGGGCTCAGAGAGACGCCCC | 517 | 1538 | AGCACAGATGCTTACCTTACCTTGGATCCCAACGTCAACCGGTGTCTCTGGGACCCATCCCTT | 1597 |
| Db | 509 | AGGGCCATGAACTCAGCTGGGAGCGGCCACACAGATGGGGCTCAGAGAGACGCCCC | 568 | 1589 | AGCACAGATGCTTACCTTGGATCCCAACGTCAACCGGTGTCTCTGGGACCCATCCCTT | 1648 |
| QY | 518 | TGCTCCAGGCGCCCGGGGCGCGCACAGAGCTGAGGGTCAACTTTGTGGCAGGTGCCG | 577 | 1598 | TTGGCATCATCTATTTGGAGCTGGCTGGCAACCACTTGAGCTTCTCTCAACTCTCTCA | 1657 |
| Db | 569 | TGCTCCAGGCGCCCGGGGCGCGCACAGAGCTGAGGGTCAACTTTGTGGCAGGTGCCG | 628 | 1649 | TTGGCATCATCTATTTGGAGCTGGCTGGCAACCACTTGAGCTTCTCTCAACTCTCTCA | 1708 |
| QY | 578 | TGAGGCCCAACAAAGGCCCTTGCCCTAGAGCGCTGTCTTGGAGGGCTGCGCGGCTTCC | 637 | 1658 | AGATGAAGATGTCCGTCATCTCTGGCGTGTGTCACATGCGCTTTGGGGTGGTCTCTCGGAG | 1717 |
| Db | 629 | TGAGGCCCAACAAAGGCCCTTGCCCTAGAGCGCTGTCTTGGAGGGCTGCGCGGCTTCC | 688 | 1709 | AGATGAAGATGTCCGTCATCTCTGGCGTGTGTCACATGCGCTTTGGGGTGGTCTCTCGGAG | 1768 |
| QY | 638 | TCAATTGCCAGCTTCAGGAGCTGGAGAGCGCGCTGGAGCACCCCGTGAACGGGCGAGCCAG | 697 | 1718 | TCTTCAACACAGTGCATCTTTGGCCAGAGGACACGGCTGTCTGTGGAGACGCTGCCCGGAGC | 1777 |
| Db | 689 | TCAATTGCCAGCTTCAGGAGCTGGAGAGCGCGCTGGAGCACCCCGTGAACGGGCGAGCCAG | 748 | 1769 | TCTTCAACACAGTGCATCTTTGGCCAGAGGACACGGCTGTCTGTGGAGACGCTGCCCGGAGC | 1828 |
| QY | 698 | CCAGTGGATGACCTTCCATCTCTACTTGGGGTCAGCAGATCGGACAGAGATCCGCA | 757 | 1778 | TCACCTTCTCTGTGGAGCTTCTCGGTTACCTCGTGTTCCTAGTCACTCAAGTGGCTGT | 1837 |
| Db | 749 | CCAGTGGATGACCTTCCATCTCTACTTGGGGTCAGCAGATCGGACAGAGATCCGCA | 808 | 1829 | TCACCTTCTCTGTGGAGCTTCTCGGTTACCTCGTGTTCCTAGTCACTCAAGTGGCTGT | 1888 |
| QY | 758 | AGATCAGGACTGCTTCCATGTCACGTCTTCCCGTTTCTGACAGCAGGAGGCCCGCC | 817 | 1838 | GTGCTGGCTGCCAGGGCGCGCTCGG---CCAGCATCTCATCTCACTTCACTCAACATGT | 1894 |
| Db | 809 | AGATCAGGACTGCTTCCATGTCACGTCTTCCCGTTTCTGACAGCAGGAGGCCCGCC | 868 | 1889 | GTGCTGGCTGCCAGGGCGCGCTCGGCCCCCAGCATCTCATCTCACTTCACTCAACATGT | 1948 |
| QY | 818 | TGCGGGCCCTGCAGCAGCTGCAACAGCAGAGCCAGAGCTGCAGAGAGTCTCTCGGGAGA | 877 | 1895 | TCTCTTCTTCCACAGCCACAGAGCTGTCTTACCCCCCGGACAGAGTGGTCCAGG | 1954 |
| Db | 869 | TGCGGGCCCTGCAGCAGCTGCAACAGCAGAGCCAGAGCTGCAGAGAGTCTCTCGGGAGA | 928 | 1949 | TCTCTTCTTCCACAGCCACAGAGCTGTCTTACCCCCCGGACAGAGTGGTCCAGG | 2008 |
| QY | 878 | CAGAGCGGTTCTTAGCCAGGTGCTAGGCGGGTGTGAGCTGTGCTGCGCCACAGGCGCAG | 937 | 1955 | CCAGCTGTGTGTCTTGGCTTGGCCATGTCGTCATCTGTCTGTGTCGACACCCCTGC | 2014 |
| Db | 929 | CAGAGCGGTTCTTAGCCAGGTGCTAGGCGGGTGTGAGCTGTGCTGCGCCACAGGCGCAG | 988 | 2009 | CCACGCTGTGTGTCTTGGCTTGGCCATGTCGTCATCTGTCTGTGTCGACACCCCTGC | 2068 |
| QY | 938 | TGAGGTCCACAAAGTGAAGGCGTGTACTGCGCCCTGAACCAAGTGCAGCGTGAACCA | 997 | 2015 | ACCTGTGTGACCGCCACCGCCCGCTGCGGAGGAGGCCCGCTGACCAACAGGAGGAAA | 2074 |
| Db | 989 | TGAGGTCCACAAAGTGAAGGCGTGTACTGCGCCCTGAACCAAGTGCAGCGTGAACCA | 1048 | 2069 | ACCTGTGTGACCGCCACCGCCCGCTGCGGAGGAGGCCCGCTGACCAACAGGAGGAAA | 2128 |
| QY | 998 | CGCACAGTGCCTCATTTGCGGAGGCTGTGTCTGTGTGAGNACTGCGCGCCCTGCAGG | 1057 | 2075 | ACAAAGCCGGGTGTGAGNACTGCGCTGAGCATCTGTGAATGGCTGGAGCTCCGATGAGG | 2134 |
| Db | 1049 | CGCACAGTGCCTCATTTGCGGAGGCTGTGTGTGTGTGTGAGNACTGCGCGCCCTGCAGG | 1108 | 2129 | ACAAAGCCGGGTGTGAGNACTGCGCTGAGCATCTGTGAATGGCTGGAGCTCCGATGAGG | 2188 |
| QY | 1058 | AGGCTCTGCGGAGACAGCTCGATGGAGGAGGAGTGTGCGGTGTCAACCGCATCCCTT | 1117 | 2135 | AAAAGGCGAGGGGCTTGGATGATGAAGAGGAGGCCGAGCTCGTCCCTCCGAGGTGCTCA | 2194 |
| Db | 1109 | AGGCTCTGCGGAGACAGCTCGATGGAGGAGGAGTGTGCGGTGTCAACCGCATCCCTT | 1168 | 2189 | AAAAGGCGAGGGGCTTGGATGATGAAGAGGAGGCCGAGCTCGTCCCTCCGAGGTGCTCA | 2248 |
| QY | 1118 | GCGGGACATGCCCCCACAATCATCCGACCAACCGCTTTCAGCGGCAGCTTTCAGGGGCA | 1177 | 2195 | TGCAACAGGCCATCCACCACTTCGAGTTCCTGCTGGGTGCGTGTCTCAACACCGCTCTCT | 2254 |
| Db | 1169 | GCGGGACATGCCCCCACAATCATCCGACCAACCGCTTTCAGCGGCAGCTTTCAGGGGCA | 1228 | 2249 | TGCACAGGCCATCCACCACTTCGAGTTCCTGCTGGGTGCGTGTCTCAACACCGCTCTCT | 2308 |
| QY | 1178 | TGTGTATCGCTACGGCTGGGCGGCTACAGAGGTCAACCCCGCTTCCATCACCATCA | 1237 | 2255 | ACCTGGCTGTGGGCGCTTGGAGCTTGGCCCAACCGCAGCTGCGAGGTTCTGTGGGGCCA | 2314 |
| Db | 1229 | TGTGTATCGCTACGGCTGGGCGGCTACAGAGGTCAACCCCGCTTCCATCACCATCA | 1288 | 2309 | ACCTGGCTGTGGGCGCTTGGAGCTTGGCCCAACCGCAGCTGTCGAGGTTCTGTGGGGCCA | 2368 |
| QY | 1238 | TCACTTCTCCCTTCTGTTGTGTGTGTGTGGGATGTGGCCACAGGGCTGTCTCATGT | 1297 | 2315 | TGCTGATGCGCATAGGCTTGGGCGCTGGGCGGAGGTGGGCGTGTGGGTGCTGG | 2374 |
| Db | 1289 | TCACTTCTCCCTTCTGTTGTGTGTGTGTGGGATGTGGCCACAGGGCTGTCTCATGT | 1348 | 2369 | TGCTGATGCGCATAGGCTTGGGCGCTTGGGCGGAGAGGTGGGCGTGTGGGTGCTGG | 2428 |
| QY | 1298 | TCTCTTCCCTTCCCTTCTGCGGAGAACCGACCGGCTGTGAAGAGCGCGCAGA | 1357 | 2375 | TCCCATCTTTCCTGCGCTTTCGCTGTATGACCGTGGCTATCTCTGTGGTGTGAGGGGAC | 2434 |
| Db | 1349 | TCTCTTCCCTTCCCTTCTGCGGAGAACCGACCGGCTGTGAAGAGCGCGCAGA | 1408 | 2429 | TCCCATCTTTCCTGCGCTTTCGCTGTATGACCGTGGCTATCTCTGTGGTGTGAGGGGAC | 2488 |
| QY | 1358 | ACGAGATCTGGCAGACTTCTTCAGGGCGGCTACCTGTCTGTCTATGGGCTGTCT | 1417 | 2435 | TCTCAGCCTTCTGTCAGCGCTTGGCTGTGCACTGGGTGGAATTCAGAACAGATTTCTACT | 2494 |
| Db | 1409 | ACGAGATCTGGCAGACTTCTTCAGGGCGGCTACCTGTCTGTCTATGGGCTGTCT | 1468 | 2489 | TCTCAGCCTTCTGTCAGCGCTTGGCTGTGCACTGGGTGGAATTCAGAACAGATTTCTACT | 2548 |
| | | | | 2495 | CAGGCACGGGCTTACAGCTGAGTCCCTTCACTTGGCTGCCACAGATGACTAGGGGCCAC | 2554 |

| | | | |
|----|------|--|------|
| Db | 2549 | CAGGCACGGGCTACAAAGCTGAGTCCCTTCACCTTCGCTGCCACAGATGACTAGGGCCAC | 2608 |
| Qy | 2555 | TGCAGGTCTCTGCCAGACCTCTCTTCTGACCTCTGAGGCAGGAGGAATAAAGACGGTCC | 2614 |
| Db | 2609 | TGCAGGTCTCTGCCAGACCTCTTCTTCTGACCTCTGAGGCAGGAGGAATAAAGACGGTCC | 2668 |
| Qy | 2615 | GCCCTGGCA | 2623 |
| Db | 2669 | GCCCTGGCA | 2677 |

Search completed: June 30, 2006, 06:44:22
Job time : 1631 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2006, 05:44:05 ; Search time 12282 Seconds
(without alignments)
12019.788 Million cell updates/sec

Title: US-10-783-519-1

Perfect score: 2640

Sequence: 1 cggcgcgcgagcgagcag.....gcaaaaaaaaaaaaaa 2640

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 48236798 seqs, 27959665780 residues

Word size : 1

Total number of hits satisfying chosen parameters: 96473154

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1: gb_est1:.*
2: gb_est3:.*
3: gb_est4:.*
4: gb_est5:.*
5: gb_est6:.*
6: gb_est7:.*
7: gb_est8:.*
8: gb_est9:.*
9: gb_est10:.*
10: gb_est11:.*
11: gb_est12:.*
12: gb_est13:.*
13: gb_est14:.*
14: gb_est15:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 1393 | 52.8 | 1864 | 6 | CR594613 full-leng |
| 2 | 1290 | 48.9 | 2493 | 14 | AY407799 Homo sapi |
| 3 | 838 | 31.7 | 1655 | 6 | CR602058 full-leng |
| 4 | 734 | 27.8 | 909 | 4 | EX358008 BX358008 |
| 5 | 729 | 27.6 | 757 | 9 | DN996103 TC110144 |
| 6 | 692 | 26.2 | 1070 | 4 | EX332189 BX332189 |
| 7 | 671 | 25.4 | 671 | 6 | CR599870 full-leng |
| 8 | 671 | 25.4 | 1088 | 3 | BQ073186 AGENCOURT |
| 9 | 620 | 23.5 | 664 | 3 | BU732254 UI-E-CQ1 |
| 10 | 619 | 23.4 | 619 | 4 | CA425567 UI-H-DF0- |
| 11 | 613 | 23.2 | 842 | 8 | CR992406 CR992406 |
| 12 | 603 | 22.8 | 603 | 5 | CD722370 OJ08h08.y |
| 13 | 602 | 22.8 | 699 | 8 | CR986883 CR986883 |
| 14 | 599 | 22.7 | 618 | 3 | BU630675 UI-H-FLO- |
| 15 | 599 | 22.7 | 619 | 3 | BU631560 UI-H-FLO- |
| 16 | 599 | 22.7 | 693 | 3 | BM976782 UI-CF-EN1 |
| 17 | 599 | 22.7 | 703 | 4 | CA307407 UI-H-FT1- |
| 18 | 599 | 22.7 | 756 | 4 | CA944661 UI-CF-FNO |
| 19 | 586 | 22.2 | 954 | 3 | BQ707846 AGENCOURT |

| | | | | | | |
|----|-----|------|------|---|----------|--------------------|
| 20 | 583 | 22.1 | 941 | 3 | BQ950273 | BQ950273 AGENCOURT |
| 21 | 583 | 22.1 | 1022 | 3 | BQ708864 | BQ708864 AGENCOURT |
| 22 | 579 | 21.9 | 764 | 4 | CA314014 | CA314014 UI-CF-FNO |
| 23 | 573 | 21.7 | 617 | 4 | CB529961 | CB529961 UI-H-FT2- |
| 24 | 573 | 21.7 | 618 | 4 | CA306854 | CA306854 UI-H-FT1- |
| 25 | 569 | 21.6 | 725 | 3 | BM977752 | BM977752 UI-CF-EN1 |
| 26 | 541 | 20.5 | 918 | 5 | CD243424 | CD243424 AGENCOURT |
| 27 | 539 | 20.4 | 590 | 2 | BI906441 | BI906441 603063470 |
| 28 | 539 | 20.4 | 634 | 4 | EX387844 | EX387844 BX387844 |
| 29 | 532 | 20.2 | 671 | 4 | EX355948 | EX355948 BX355948 |
| 30 | 529 | 20.0 | 641 | 4 | EX355947 | EX355947 BX355947 |
| 31 | 529 | 20.0 | 803 | 7 | BF690458 | BF690458 602186921 |
| 32 | 528 | 20.0 | 911 | 4 | EX328356 | EX328356 BX328356 |
| 33 | 526 | 19.9 | 921 | 4 | EX358007 | EX358007 BX358007 |
| 34 | 521 | 19.7 | 574 | 7 | AW118573 | AW118573 xd94b04.x |
| 35 | 521 | 19.7 | 1010 | 3 | BQ059934 | BQ059934 AGENCOURT |
| 36 | 521 | 19.7 | 1011 | 4 | EX384133 | EX384133 BX384133 |
| 37 | 519 | 19.7 | 916 | 4 | EX372358 | EX372358 BX372358 |
| 38 | 515 | 19.5 | 781 | 2 | EG422109 | EG422109 602448826 |
| 39 | 514 | 19.5 | 679 | 8 | CN420054 | CN420054 170005325 |
| 40 | 512 | 19.4 | 520 | 8 | CN420053 | CN420053 170006000 |
| 41 | 512 | 19.4 | 879 | 2 | BI518962 | BI518962 603061885 |
| 42 | 503 | 19.1 | 916 | 2 | EG121525 | EG121525 602352860 |
| 43 | 500 | 18.9 | 898 | 4 | CA454750 | CA454750 AGENCOURT |
| 44 | 499 | 18.9 | 906 | 3 | BU855490 | BU855490 AGENCOURT |
| 45 | 498 | 18.9 | 1074 | 4 | EX385318 | EX385318 BX385318 |

ALIGNMENTS

CR594613 1864 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CSDD004YI18 of Neuroblastoma Cot
50-normalized of Homo sapiens (human).

CR594613.1 GI:50475420

HTC; CNSLT_cDNA.

Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

1 (bases 1 to 1864)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Feng Liang Email: fliang@lifetech.com URL:

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue

2 (bases 1 to 1864)

Genoscope.

Direct Submission

Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: seque@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

Location/Qualifiers

1. 1864

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSDD004YI18"

/tissue_type="Neuroblastoma Cot 50-normalized"

/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 52.8%; Score 1393; DB 6; Length 1864;

Best Local Similarity 99.6%; Pred. No. 0;
Matches 1833; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

```
QY 770 GCTTCCACTGCGACGCTGCTCCCGTTCTGTCAGACAGAGAGGCGCCGCTCGGGGCGCTGC 829
Db 25 GCTTCCACTGCGACGCTGCTCCCGTTCTGTCAGACAGAGAGGCGCCGCTCGGGGCGCTGC 84
QY 830 AGCAGCTGCAACAGCAGACAGAGAGCTGCGAGAGGTCTCCGCGGAGAGACAGAGCGGTCC 889
Db 85 AGCAGCTGCAACAGCAGACAGAGAGCTGCGAGAGGTCTCCGCGGAGAGACAGAGCGGTCC 144
QY 890 TGAGCCAGGTGCTAGGCGCGGCTGCTGACGCTGCTGCGCCAGGCGAGGTGCGAGTCCACA 949
Db 145 TGAGCCAGGTGCTAGGCGCGGCTGCTGACGCTGCTGCGCCAGGCGAGGTGCGAGTCCACA 204
QY 950 AGATGAAGCCGCTGTACTGCGCCCTGAAACAGTGCAGAGCTGAGACACAGCAAGTGCC 1009
Db 205 AGATGAAGCCGCTGTACTGCGCCCTGAAACAGTGCAGAGCTGAGACACAGCAAGTGCC 264
QY 1010 TCATTGCCGAGGCTGCTGCTGTCGAGACCTGCGCCCTGACAGGAGGCGCTCGGG 1069
Db 265 TCATTGCCGAGGCTGCTGCTGTCGAGACCTGCGCCCTGACAGGAGGCGCTCGGG 324
QY 1070 ACAGCTCGATGAGGAGGAGTGAAGTCCGCTGCTCACCGCATCCGCTGCGCGGACATGC 1129
Db 325 ACAGCTCGATGAGGAGGAGTGAAGTCCGCTGCTCACCGCATCCGCTGCGCGGACATGC 384
QY 1130 CCGCCACATCATCCGCAACACCGCTTACGCGCAGCTTCCAGGCGATCGTGGATCGCT 1189
Db 385 CCGCCACATCATCCGCAACACCGCTTACGCGCAGCTTCCAGGCGATCGTGGATCGCT 444
QY 1190 ACGGCTGCGCGCTACACAGGAGTCAACCCGCTCCCTACACATCATCATCCCTCCCT 1249
Db 445 ACGGCTGCGCGCTACACAGGAGTCAACCCGCTCCCTACACATCATCATCCCTCCCT 504
QY 1250 TCTGTTTGTGTGATGTTTGGGGATGTGGGCAACCGGCTGCTCATGTTCTTCCGCC 1309
Db 505 TCTGTTTGTGTGATGTTTGGGGATGTGGGCAACCGGCTGCTCATGTTCTTCCGCC 564
QY 1310 TGGCCATGCTCTTGGGAGAACCGACCGCTGTGAAGCGCGCGACAGATCTGGC 1369
Db 565 TGGCCATGCTCTTGGGAGAACCGACCGCTGTGAAGCGCGCGACAGATCTGGC 624
QY 1370 AGACTTTCTTACAGGCGCGCTACTGCTCTGCTATGAGCGCTGTTCTCATCTACACCG 1429
Db 625 AGACTTTCTTACAGGCGCGCTACTGCTCTGCTATGAGCGCTGTTCTCATCTACACCG 684
QY 1430 GCTTCATCTACAAACGAGTGTTCAGTGCGCCACACAGCATCTTCCGCTCGGGGTGGAGTG 1489
Db 685 GCTTCATCTACAAACGAGTGTTCAGTGCGCCACACAGCATCTTCCGCTCGGGGTGGAGTG 744
QY 1490 TGGCGGCCATGGCCAAACGATCTGGCTGAGTGATGATCTTCTGGCCGACGACAGATGC 1549
Db 745 TGGCGGCCATGGCCAAACGATCTGGCTGAGTGATGATCTTCTGGCCGACGACAGATGC 804
QY 1550 TTACCTCGATCCCAACGTCACCGGTGCTTCTGCGGACCGCTACCCCTTTGGCATCGATC 1609
Db 805 TTACCTCGATCCCAACGTCACCGGTGCTTCTGCGGACCGCTACCCCTTTGGCATCGATC 864
QY 1610 CTATTTGGAGCCTGGCTGCCAACCACTTGAGCTTCTCAACTCTCTTCAAGATGAAGATGT 1669
Db 865 CTATTTGGAGCCTGGCTGCCAACCACTTGAGCTTCTCAACTCTCTTCAAGATGAAGATGT 924
QY 1670 CGGTATCTTGGGCTGCTGTCATGCGCTTTGGGGTGTGCTCGAGTCTTCAACACG 1729
Db 925 CGGTATCTTGGGCTGCTGTCATGCGCTTTGGGGTGTGCTCGAGTCTTCAACACG 984
QY 1730 TGCACCTTTCGCGCAGAGGACCGCTGCTGTCGAGACGCTGCGGAGCTCACCTTCTCTGC 1789
Db 985 TGCACCTTTCGCGCAGAGGACCGCTGCTGTCGAGACGCTGCGGAGCTCACCTTCTCTGC 1044
QY 1790 TGGGACTCTTCGCTTACCTCGTGTCTTCTAGTATCATCTACAGTGGCTGTGTCTGGCTG 1849
```

```
Db 1045 TGGGACTCTTCGTTACCTCGTGTTCCTAGTATCATCTACAAGTGGCTGTGTGTGGCTG 1104
QY 1850 CAGAGGCGCGCTCG---CCAGCATCTCTCATCTCACTTCAATCAACATGTCTCTTCTCCC 1906
Db 1105 CAGAGGCGCGCTCGGCCCCCAGCATCTCTCACTTCAATCAACATGTCTCTTCTCCC 1164
QY 1907 ACAGCCCCCAGCAACAGGCTGCTTACCCCGGAGAGGTGTCCAGAGCCACGCTGGTGG 1966
Db 1165 ACAGCCCCCAGCAACAGGCTGCTTACCCCGGAGAGGTGTCCAGGCCACGCTGGTGG 1224
QY 1967 TCTGCGCTTGGCCATGTGTGCTGCTGCTTGGCACACCCCTGCACTGCTGCTGCAACC 2026
Db 1225 TCTGCGCTTGGCCATGTGTGCTGCTGCTTGGCACACCCCTGCACTGCTGCTGCAACC 1284
QY 2027 GGCACCGCGCGCTGCGGAGGAGGCGCGCTGACCGACAGAGAGGAAACAGAGCCCGGT 2086
Db 1285 GGCACCGCGCGCTGCGGAGGAGGCGCGCTGACCGACAGAGAGGAAACAGAGCCCGGT 1344
QY 2087 TCTGCGACCTGCTGACGCGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGAGGAGGG 2146
Db 1345 TCTGCGACCTGCTGACGCGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGAGGAGGG 1404
QY 2147 GCCTGATGATGAAGAGGAGGCGGAGCTGCTGCCCTCCGAGGTGCTCATGCAACAGGCCA 2206
Db 1405 GCCTGATGATGAAGAGGAGGCGGAGCTGCTGCCCTCCGAGGTGCTCATGCAACAGGCCA 1464
QY 2207 TCACACCATCGAGTTCTGCTGGGCTGCTTCCAAACCCGCTCTCTACCTGCGGCTGT 2266
Db 1465 TCACACCATCGAGTTCTGCTGGGCTGCTTCCAAACCCGCTCTCTACCTGCGGCTGT 1524
QY 2267 GGGCCCTGAGCTGCGCCACCGCAGCTGTCGAGGTTCTGCGGCAATGATGATGCGCA 2326
Db 1525 GGGCCCTGAGCTGCGCCACCGCAGCTGTCGAGGTTCTGCGGCAATGATGATGCGCA 1584
QY 2327 TAGGCTTGGGCTGCGCGGAGGTGGGCTGGCGCTGTGTGTGCTGCTGCCCATCTTTG 2386
Db 1585 TAGGCTTGGGCTGCGCGGAGGTGGGCTGGCGCTGTGTGTGCTGCTGCCCATCTTTG 1644
QY 2387 CGGCTTTGCGGTGATGACCGTGGCTATCTGCTGGTGTGATGAGGACCTCTAGCCTTCC 2446
Db 1645 CGGCTTTGCGGTGATGACCGTGGCTATCTGCTGGTGTGATGAGGACCTCTAGCCTTCC 1704
QY 2447 TCACGCGCTGCGGCTGCACTGGGTGGAATTCAGAACAAAGTCTTACTCAGGACCGGCT 2506
Db 1705 TCACGCGCTGCGGCTGCACTGGGTGGAATTCAGAACAAAGTCTTACTCAGGACCGGCT 1764
QY 2507 ACAAGCTGAGTCCCTTCACTTCGCTGCCACAGATGACTAGGCGCCACTGCGAGTCTCTG 2566
Db 1765 ACAAGCTGAGTCCCTTCACTTCGCTGCCACAGATGACTAGGCGCCACTGCGAGTCTCTG 1824
QY 2567 CAGAGCTCTTCTGACCTCTGAGGCGAGGAGGAATAAA 2606
Db 1825 CAGAGCTCTTCTGACCTCTGAGGCGAGGAGGAATAAA 1864
```

RESULT 2

AY407799

LOCUS

DEFINITION

Homo sapiens

AY407799

ACCESSION

AY407799

VERSION

AY407799.1

KEYWORDS

GSS.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

AY407799 2493 bp DNA linear GSS 15-DEC-2003
Homo sapiens TCIRG1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY407799.1 GI:39763770
GSS.
Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1. (bases 1 to 2493)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.

TITLE
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2493)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source Location/Qualifiers
1. .2493
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
gene <1..>2493
/gene="TCIRG1"
/locus_tag="HCM3011"
ORIGIN
Query Match 48.9%; Score 1290; DB 14; Length 2493;
Best Local Similarity 99.6%; Pred. No. 0; Mismatches 3; Gaps 1;
Matches 1680; Conservative 0; Indels 3;
1;
QY 865 GTCTCTGGGAGACAGACGGCTTCTGAGCCAGTGTAGCCGGGTGTCAGCTGCTG 924
DB 808 GTCTCTGGGAGACAGACGGCTTCTGAGCCAGTGTAGCCGGGTGTCAGCTGCTG 867
QY 925 CCGCAGGGCAGGTGCAAGTCCACAAGATGAAGCCCTGTACTTGGCCCTGAACAGTGC 984
DB 868 CCGCAGGGCAGGTGCAAGTCCACAAGATGAAGCCCTGTACTTGGCCCTGAACAGTGC 927
QY 985 AGGTGAGCACCAGCAGAGTGCCTCATTTGCCAGCCCTGCTCTGTGCGAGCTG 1044
DB 928 AGGTGAGCACCAGCAGAGTGCCTCATTTGCCAGCCCTGCTCTGTGCGAGCTG 987
QY 1045 CCGCCCTCGAGGAGCCCTCGGGACAGCTCGATGGAGGAGGTGAGTGCCTGGCT 1104
DB 988 CCGCCCTCGAGGAGCCCTCGGGACAGCTCGATGGAGGAGGTGAGTGCCTGGCT 1047
QY 1105 CACCGCATCCCTGCGGGACATGCCCCCACTCATTCGCAACCAACCCGCTTACCGCC 1164
DB 1048 CACCGCATCCCTGCGGGACATGCCCCCACTCATTCGCAACCAACCCGCTTACCGCC 1107
QY 1165 AGCTTCCAGGGCATCGTGGATCGCTAGCGGTGGCCGCTACAGGAGGTCAACCCGCT 1224
DB 1108 AGCTTCCAGGGCATCGTGGATCGCTAGCGGTGGCCGCTACAGGAGGTCAACCCGCT 1167
QY 1225 CCCTACACCATCATCATTCCCTTCTCTGTGCTGTGATGTTGCGGGATGTGGGCGAC 1284
DB 1168 CCCTACACCATCATCATTCCCTTCTCTGTGCTGTGATGTTGCGGGATGTGGGCGAC 1227
QY 1285 GGGGTGCTCATGTTCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 1344
DB 1228 GGGGTGCTCATGTTCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT 1287
QY 1345 AAGCCCGCAGAACAGATCTGGACACTTTCTTTCAGGGCCGCTACCTGCTTCTGCTT 1404
DB 1288 AAGCCCGCAGAACAGATCTGGACACTTTCTTTCAGGGCCGCTACCTGCTTCTGCTT 1347
QY 1405 ATGGCCCTGTTCTTCATCTACCGGCTTCTATCTACAAAGAGTCTTTCAGTCCGCCACC 1464
DB 1348 ATGGCCCTGTTCTTCATCTACCGGCTTCTATCTACAAAGAGTCTTTCAGTCCGCCACC 1407
QY 1465 AGCATCTTCCCTCGGGCTGGAGTGTGGCCGCTACCGCAACCAAGTCTGCTGGAGTGT 1524
DB 1408 AGCATCTTCCCTCGGGCTGGAGTGTGGCCGCTACCGCAACCAAGTCTGCTGGAGTGT 1467
QY 1525 GCATTCTGCCCCAGCACAGATGCTTACCTGGATCCCAAGCTACCGGTGCTTCTCTG 1584

Db 1468 GCATTCTGCCCCAGCACAGATGCTTACCTTGGATCCCAACGTACCGGTGCTTCTCTG 1527
QY 1585 GGACCTTACCCCTTTTGGCATCGATCTTATTTGGAGCCCTGGCTGCAACCACTTGAAGTTC 1644
DB 1528 GGACCTTACCCCTTTTGGCATCGATCTTATTTGGAGCCCTGGCTGCAACCACTTGAAGTTC 1587
QY 1645 CTCAACTCTCTTCAAGATGAAGATGTCGTCATCTTGGGCGTGTGTCATGTCATGCTTGGG 1704
DB 1588 CTCAACTCTCTTCAAGATGAAGATGTCGTCATCTTGGGCGTGTGTCATGTCATGCTTGGG 1647
QY 1705 GTGTCTCTCGAGTCTTCAACACAGTCAGTCTTTCGCGCAGAGGACCGCTGCTGTCGAG 1764
DB 1648 GTGTCTCTCGAGTCTTCAACACAGTCAGTCTTTCGCGCAGAGGACCGCTGCTGTCGAG 1707
QY 1765 AGCTGTCGGAGCTCACCTTCTCTGCTGGGACTCTTTCGGTTCACCTGCTGTTCTTCTAGTCA 1824
DB 1708 AGCTGTCGGAGCTCACCTTCTCTGCTGGGACTCTTTCGGTTCACCTGCTGTTCTTCTAGTCA 1767
QY 1825 TACAAGTGGCTGTGTGTCGGGTCGAGGGCCGCTCG---CCAGCATCTCTATCCAC 1881
DB 1768 TACAAGTGGCTGTGTGTCGGGTCGAGGGCCGCTCGGGCCCGCAGCATCTCTATCCAC 1827
QY 1882 TTCAATCAACATGTTCTTCTCCACAGCCCAAGCAACAGGCTGCTTACCCCGGCGAG 1941
DB 1828 TTCAATCAACATGTTCTTCTCCACAGCCCAAGCAACAGGCTGCTTACCCCGGCGAG 1887
QY 1942 GAGTGTGTCAGGCGCACGCTGCTGCTTTCGCTTGGCCATGTTGTCCTCTCTGCTT 2001
DB 1888 GAGTGTGTCAGGCGCACGCTGCTGCTTTCGCTTGGCCATGTTGTCCTCTCTGCTT 1947
QY 2002 GGCACACCTCTGACCTGTGTGACCGCCACCGCCGCTGTCGGAGAGGAGGCGCTGAC 2061
DB 1948 GGCACACCTCTGACCTGTGTGACCGCCACCGCCGCTGTCGGAGAGGAGGCGCTGAC 2007
QY 2062 CGACAGAGAGAAACAAGGCGGCTGCTGACCTGCTGACGATCTGTGAATGGCTGG 2121
DB 2008 CGACAGAGAGAAACAAGGCGGCTGCTGACCTGCTGACGATCTGTGAATGGCTGG 2067
QY 2122 AGCTCCGATGAGGAAAGGCGGCTGATGATGAAGAGGAGGCGGCTGCTGCTCC 2181
DB 2068 AGCTCCGATGAGGAAAGGCGGCTGATGATGAAGAGGAGGCGGCTGCTGCTCC 2127
QY 2182 TCGAGGTGCTCATGACACAGGCGCATCCACACATCGAGTTCCTGCTGGGCTGCTGCTCC 2241
DB 2128 TCGAGGTGCTCATGACACAGGCGCATCCACACATCGAGTTCCTGCTGGGCTGCTGCTCC 2187
QY 2242 AACACCGCTTCTACCTGCGGCTGTCGCGCTGAGCCCTGAGCCACGCGCTGCTGCGAG 2301
DB 2188 AACACCGCTTCTACCTGCGGCTGTCGCGCTGAGCCCTGAGCCACGCGCTGCTGCGAG 2247
QY 2302 GTTCTGTGGGCGCATGTCGCGATAGGCGCTGCGGCTGCGGCTGCGGAGGTGGGCG 2361
DB 2248 GTTCTGTGGGCGCATGTCGCGATAGGCGCTGCGGCTGCGGCTGCGGAGGTGGGCG 2307
QY 2362 GCTGTGTGCTGTCCTTCCATCTTTCGCGCTTTCGCGCTGATGACCGTGGCTATCTGCTG 2421
DB 2308 GCTGTGTGCTGTCCTTCCATCTTTCGCGCTTTCGCGCTGATGACCGTGGCTATCTGCTG 2367
QY 2422 GTGATGAGGAGTCTCAGCTTTCGCGCTTTCGCGCTGAGTCCCTTTCGCTGCGGATTCAG 2481
DB 2368 GTGATGAGGAGTCTCAGCTTTCGCGCTTTCGCGCTGAGTCCCTTTCGCTGCGGATTCAG 2427
QY 2482 AACAAAGTTCCTTACCTCAGGCGCGGTACAAGCTGAGTCCCTTTCGCTGCGGATTCAG 2541
DB 2428 AACAAAGTTCCTTACCTCAGGCGCGGTACAAGCTGAGTCCCTTTCGCTGCGGATTCAG 2487
QY 2542 GACTAG 2547
DB 2488 GACTAG 2493

RESULT 3

```
CR602058      1655 bp      mRNA      linear      HTC 21-JUL-2004
LOCUS      full-length cDNA clone CS0DI059YF06 of Placenta Cot 25-normalized
DEFINITION      of Homo sapiens (human).
ACCESSION      CR602058
VERSION      CR602058.1 GI:50482865
KEYWORDS      HTC; CNSLT CDNA.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominiidae; Homo.
REFERENCE      1 (bases 1 to 1655)
AUTHORS      Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished
REMARK      Contact : Feng Liang Email : fliang@lifetech.com URL :
              http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
              Paraday Avenue
              2 (bases 1 to 1655)
REFERENCE      Genoscope.
AUTHORS      Direct Submission
TITLE      Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL      BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT      1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
              end enriched, double-strand cDNA was digested with Not I and cloned
              into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
              was normalized. Library was constructed by Life Technologies, a
              division of Invitrogen.
FEATURES
              Location/Qualifiers
              1..1655
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="CS0DI059YF06"
               /tissue type="Placenta Cot 25-normalized"
               /plasmid="pCMVSPORT_6"
ORIGIN
Query Match      31.7%; Score 838; DB 6; Length 1655;
Best Local Similarity 99.7%; Pred. NO. 0;
Matches 1078; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY      1519      AGTGATGCAATTCCTGGCCAGCACGATGCTTACCTGGATCCCAAGCTCACCGGTGC 1578
DB      575      AGTGATGCAATTCCTGGCCAGCACGATGCTTACCTGGATCCCAAGCTCACCGGTGC 634

QY      1579      TTCTGGGACCTTACCCCTTTGGCATCGATCCTATTGTGGAGCTGGCTGCCAACCACTTG 1638
DB      635      TTCTGGGACCTTACCCCTTTGGCATCGATCCTATTGTGGAGCTGGCTGCCAACCACTTG 694

QY      1639      AGCTTCTCAACTCTTCAAGATGAAGATGTCGTCATCCTGGGCGTGTGTGACATGGCC 1698
DB      695      AGCTTCTCAACTCTTCAAGATGAAGATGTCGTCATCCTGGGCGTGTGTGACATGGCC 754

QY      1699      TTGTGGGTGGTCTCGGAGCTTCAACACGTCGCTTGGCCAGAGGACCGGCTGCTG 1758
DB      755      TTGTGGGTGGTCTCGGAGCTTCAACACGTCGCTTGGCCAGAGGACCGGCTGCTG 814

QY      1759      CTGGAGACGCTCGGAGCTCACCTTCTGCTGGGACTCTTGGTTACCTCGTGTTCCTA 1818
DB      815      CTGGAGACGCTCGGAGCTCACCTTCTGCTGGGACTCTTGGTTACCTCGTGTTCCTA 874

QY      1819      GTCATCTCAAGTGGTGTGTGTGGGCTGCGAGGGCGGCTCG---CCGAGCATCTTC 1875
DB      875      GTCATCTCAAGTGGTGTGTGTGGGCTGCGAGGGCGGCTCGGAGGATCTCTC 934

QY      1876      ATCCACTTCAACATGTTCTTCTTCTCCACAGCCCGCAGCAACAGGCTGCTTACCCC 1935
DB      935      ATCCACTTCAACATGTTCTTCTTCTCCACAGCCCGCAGCAACAGGCTGCTTACCCC 994

QY      1936      CGGCAGGAGGTGGTCCAGGCCACGCTGGTGGTCTGGCCCTTGGCCATGGTGGCCATCCTG 1995
```

```
Db      995      CGCAGGAGGTGGTCCAGGCCACGCTGGTGGTCTTGGCCATGGTGGCCATCCTG 1054
QY      1996      CTGCTTGGACACACCTCTGCACTGTGTGACACCGCCACCGCCGCGCTGCGAGGAGGCC 2055
DB      1055      CTGCTTGGACACACCTCTGCACTGTGTGACACCGCCACCGCCGCGCTGCGAGGAGGCC 1114
QY      2056      GCTGACCCGACAGGAGGAAACAAAGCCGGTGTGTGACCTGCTGACGACATCTGTAAT 2115
DB      1115      GCTGACCCGACAGGAGGAAACAAAGCCGGTGTGTGACCTGCTGACGACATCTGTAAT 1174
QY      2116      GCTGAGACTCCGATGAGGAAAGCAGGGGCTTGGATGATGAAGAGGAGGCGCAGCTC 2175
DB      1175      GCTGAGACTCCGATGAGGAAAGCAGGGGCTTGGATGATGAAGAGGAGGCGCAGCTC 1234
QY      2176      GTCCCTCGAGGTGCTCATGACACAGGCTATCAACATCGAGTTCTGCTGGGCTGC 2235
DB      1235      GTCCCTCGAGGTGCTCATGACACAGGCTATCAACATCGAGTTCTGCTGGGCTGC 1294
QY      2236      GTCTCCACACCGCTCTTACCTGCGCTGTGGGCTTGGGCTGGCCACGCGCAGCTG 2295
DB      1295      GTCTCCACACCGCTCTTACCTGCGCTGTGGGCTTGGGCTGGCCACGCGCAGCTG 1354
QY      2296      TCCGAGGTTCTGTGGGCTATGATGCGCATAGGCTGGGCTGGGCGCGGAGGTGGGC 2355
DB      1355      TCCGAGGTTCTGTGGGCTATGATGCGCATAGGCTGGGCTGGGCGCGGAGGTGGGC 1414
QY      2356      GTGGCGGTGTGGTGTGCTGCTCCATCTTTGGCGCTTTGGCGTGAATGACGTTGCTATC 2415
DB      1415      GTGGCGGTGTGGTGTGCTGCTCCATCTTTGGCGCTTTGGCGTGAATGACGTTGCTATC 1474
QY      2416      CTGCTGGTGTGATGGAGGACTCTCAGCTTCTGCGACGCTGGGCTGGGCTGACCTGGTGA 2475
DB      1475      CTGCTGGTGTGATGGAGGACTCTCAGCTTCTGCGACGCTGGGCTGGGCTGACCTGGTGA 1534
QY      2476      TTCCAGAACAGTTCTACTCAGGACGCGGTACAAAGTGAAGTCCCTTCACTTCCGCTGC 2535
DB      1535      TTCCAGAACAGTTCTACTCAGGACGCGGTACAAAGTGAAGTCCCTTCACTTCCGCTGC 1594
QY      2536      ACAGATGATAGGGCCCACTGAGTCTGCGACGCTTCTTCCAGACCTTCTTCCAGACGAG 2595
DB      1595      ACAGATGATAGGGCCCACTGAGTCTGCGACGCTTCTTCCAGACCTTCTTCCAGACGAG 1654
QY      2596      A 2596
DB      1655      A 1655

RESULT 4
BX358008      909 bp      mRNA      linear      EST 08-APR-2004
LOCUS      BX358008 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION      clone CS0DI032YN10 5-PRIME, mRNA sequence.
ACCESSION      BX358008
VERSION      BX358008.2 GI:46285730
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominiidae; Homo.
              1 (bases 1 to 909)
              Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
              Full-length cDNA libraries and normalization
              Unpublished (2001)
              On May 5, 2003 this sequence version replaced gi:30366180.
              Contact: Genoscope
              Genoscope - Centre National de Sequencage
              2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
              Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
              1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
              end enriched, double-strand cDNA was digested with Not I and cloned
              into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
```


was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6186.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0D1032D05Q0P1&c=6186.f.

FEATURES

source
1. 909
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1032YN10"
/tissue type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 27.8%; Score 734; DB 4; Length 909;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 884; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 607 CGCTCTCTGGAGGCGCTGCGCGCTTCTCTATTGCCAGCTTCAGGAGCTGGAGCAG 666
Db 9 CGCTCTCTGGAGGCGCTGCGCGCTTCTCTATTGCCAGCTTCAGGAGCTGGAGCAG 68
QY 667 CGCTCTGGAGCACCCTGTCAGCGCGGAGCCAGCCAGCTGATGATCTCTCTCTCTCTAC 726
Db 69 CGCTCTGGAGCACCCTGTCAGCGCGGAGCCAGCCAGCTGATGATCTCTCTCTCTAC 128
QY 727 TGGGGTGGAGATCGGACAGAGATCGGCAAGATCAGGACCTGCTTCCATCCACGTC 786
Db 129 TGGGGTGGAGATCGGACAGAGATCGGCAAGATCAGGACCTGCTTCCATCCACGTC 188
QY 787 TTCCCGTTCTGCAGCAGGAGGAGCGCGCTCGGGGCCCTGCAGAGCTGCAACAGCAG 846
Db 189 TTCCCGTTCTGCAGCAGGAGGAGCGCGCTCGGGGCCCTGCAGAGCTGCAACAGCAG 248
QY 847 AGCCAGGAGCTGCAGGAGGTCTCGGGGAGACAGAGCGGTCTTGAGCCAGGTGCTAGGC 906
Db 249 AGCCAGGAGCTGCAGGAGGTCTCGGGGAGACAGAGCGGTCTTGAGCCAGGTGCTAGGC 308
QY 907 CGGCTCTGAGCTGCTCGGCCAGGCGAGGTGCAGGTCCAAAGATGAAGCCGTGTAC 966
Db 309 CGGCTCTGAGCTGCTCGGCCAGGCGAGGTGCAGGTCCAAAGATGAAGCCGTGTAC 368
QY 967 CTGGCCCTGAACAGTGCAGCTGAGCACCAGCAGAGTCTCTATTCGCGAGGCTGG 1026
Db 369 CTGGCCCTGAACAGTGCAGCTGAGCACCAGCAGAGTCTCTATTCGCGAGGCTGG 428
QY 1027 TGCTCTGTGAGACCTGCGCGCCCTGCAGGAGGCGCTCGGGACAGCTCGATGGAGGAG 1086
Db 429 TGCTCTGTGAGACCTGCGCGCCCTGCAGGAGGCGCTCGGGACAGCTCGATGGAGGAG 488
QY 1087 GAGTGAAGTCCGTGCTACCGCATCCCTGCGGGAGATGCCCCCAGCTCATCCG 1146
Db 489 GAGTGAAGTCCGTGCTACCGCATCCCTGCGGGAGATGCCCCCAGCTCATCCG 548
QY 1147 ACCAACCGCTTCAAGGCGAGCTTCCAGGCGATCGTGGATCGCTACCGGCTGGCGCTAC 1206
Db 549 ACCAACCGCTTCAAGGCGAGCTTCCAGGCGATCGTGGATCGCTACCGGCTGGCGCTAC 608
QY 1207 CAGGAGGTCAACCCCGCTCTACACCATCATCCTTCCCTTCTCTGTTGCTGTGATG 1266
Db 609 CAGGAGGTCAACCCCGCTCTACACCATCATCCTTCCCTTCTCTGTTGCTGTGATG 668
QY 1267 TTCGGGGATGTGGCCACCGGCTGTCTCATGTTCTCTTCCCTCTGGCCATGTCCTTGG 1326
Db 669 TTCGGGGATGTGGCCACCGGCTGTCTCATGTTCTCTTCCCTCTGGCCATGTCCTTGG 728
QY 1327 GAGAACCGAGCGCTGTGAAGCGCGCAGAGAGATCTGGCAGACTTCTTCAGGGGC 1386

Db 729 GAGAACCGAGCGCTGTGAAGCGCGCGCAGAACGAGATCTGGCAGACTTCTTCAGGGGC 788
QY 1387 CGCTACTGCTCTGCTCTGCTTATGGGCGCTTCTTCCATCTTACACCGGCTTCACTACACGAG 1446
Db 789 CGCTACTGCTCTGCTTATGGGCGCTTCTTCCATCTTACACCGGCTTCACTACACGAG 848
QY 1447 TGCTTACTGTCGCGCCAGCAGCATCTTCCCTCGGCGCTGGAGTGTGGC 1493
Db 849 TGCTTACTGTCGCGCCAGCAGCATCTTCCCTCGGCGCTGGAGTGTGGC 895

RESULT 5

DN996103
LOCUS
DEFINITION
Tc110144 Human breast cancer tissue, large insert, pCMV expression library Homo sapiens cDNA clone Tc110144 5', similar to Homo sapiens T-cell, immune regulator 1, ATPase, H+ transporting, lysosomal V0 protein a isoform 3 (TCIRG1), transcript variant 2, mRNA sequence.
DN996103
VERSION
DN996103.1 GI:66255930
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

ACCESSION

DN996103
VERSION
DN996103.1 GI:66255930
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 757)
AUTHORS
Birkett,C., Cho,J., Gau,Y., Hamer,R., Kelly,S., Kovacs,K., Liu,L., Liu,X., Porter,J., Sachs,A., Shu,Y., Sun,Z., Wong,J., Wu,M., Zhang,X., Jay,G. and He,W.
High-throughput cloning of full-length human cDNAs directly from cDNA libraries optimized for large and rare transcripts
Unpublished (2005)
Contact: Kovacs, KF

TITLE

High Throughput cDNA Cloning
Origene Technologies, Inc. (www.origene.com)
6 Taft Court, Suite 100, Rockville, MD 20850, USA
Tel: 301 340 3188
Fax: 301 340 8606
Email: cdna@origene.com

JOURNAL

This EST submission is part of an on-going human full-length cloning project at Origene Technologies, Inc.
Please contact Origene for access.
Origene Technologies, Inc.
6 Taft Ct. Suite 100
Rockville, MD 20850
Tel: (301) 340-3188
http://www.origene.com
Seq primer: pCMV6 5prime forward vector primer, Origene Technologies Inc.

COMMENT

High Throughput cDNA Cloning
Origene Technologies, Inc. (www.origene.com)
6 Taft Court, Suite 100, Rockville, MD 20850, USA
Tel: 301 340 3188
Fax: 301 340 8606
Email: cdna@origene.com

FEATURES

Location/Qualifiers
1. .757
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TC110144"
/tissue type="Breast cancer"
/clone_lib="Human breast cancer tissue, large insert, pCMV expression library"
/note="Organ: Mammary gland (cancer tissue); Vector: pCMV6-XL5; Site_1: EcoRI; Site_2: XhoI/Sall compatible end ligatio; Oligo-dT primed reverse transcription optimized for large and GC rich mRNA transcripts, cDNA size selection, optimized ligation for large inserts into mammalian expression vector, random clones selected for end sequence verification of full-length genes"

ORIGIN

Query Match 27.6%; Score 729; DB 9; Length 757;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 729; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 377 AGTCGGGATGTGGGGGCAACCGAGCCCTCGGGGCCAGCTGCACGAGTGCAGC 436

```
Db 29 AGCTGCGGATGTCGGGGCAACAGAGAGCCCTCGGGCCAGCTGCACCACTGAGC 88
Qy 437 TCACGCGCGCTGTACGCCAGGGCCATGAACCTCAGCTGGCAGCGCCCAACAGATG 496
Db 89 TCACGCGCGCTGTACGCCAGGGCCATGAACCTCAGCTGGCAGCGCCCAACAGATG 148
Qy 497 GGGCTTCAGAGAGAGCGCCCTGCTCCAGGCCCCCGGGGGCCGACACAGGACCTGAGG 556
Db 149 GGGCTTCAGAGAGAGCGCCCTGCTCCAGGCCCCCGGGGGCCGACACAGGACCTGAGG 208
Qy 557 TCAACTTTGTGACAGTGCCTGAGAGCCCAAGGCCCTTCCCTTAGAGCGCTGCTCT 616
Db 209 TCAACTTTGTGACAGTGCCTGAGAGCCCAAGGCCCTTCCCTTAGAGCGCTGCTCT 268
Qy 617 GAGGGCTGCGCGGCTTCTCATTCAGAGTTCAGGAGTGCAGAGCGCGCTGAGC 676
Db 269 GAGGGCTGCGCGGCTTCTCATTCAGAGTTCAGGAGTGCAGAGCGCGCTGAGC 328
Qy 677 ACCCGTTCAGCGGCGAGCGAGCCAGTGCATGACCTTCTCATCTCTACTGCGGTGAGC 736
Db 329 ACCCGTTCAGCGGCGAGCGAGCCAGTGCATGACCTTCTCATCTCTACTGCGGTGAGC 388
Qy 737 AGATCGGACAGAGATCCGCAAGATCACGAGTGTCTCCACTGCGCACGCTTCCCGTTTC 796
Db 389 AGATCGGACAGAGATCCGCAAGATCACGAGTGTCTCCACTGCGCACGCTTCCCGTTTC 448
Qy 797 TGCAGCAGGAGAGCGCGCTCTCGGGGCGCTCAGCAGTGCAGAGTGCAGAGCCAGGAGC 856
Db 449 TGCAGCAGGAGAGCGCGCTCTCGGGGCGCTCAGCAGTGCAGAGTGCAGAGCCAGGAGC 508
Qy 857 TGCAGAGTCTCTCGGGGAGACAGAGCGGTTCTCGAGCAGTGTAGCGCGGCTGTC 916
Db 509 TGCAGAGTCTCTCGGGGAGACAGAGCGGTTCTCGAGCAGTGTAGCGCGGCTGTC 568
Qy 917 AGCTGCTGCGCGCAGGCGAGTGCAGGTCACAAGATGAAGCGCTGTACCTGGCCCTGA 976
Db 569 AGCTGCTGCGCGCAGGCGAGTGCAGGTCACAAGATGAAGCGCTGTACCTGGCCCTGA 628
Qy 977 ACCAGTGCAGGTCAGCAGCAGCAGCAGTGCATTCAGGAGCGCTGCTGTGTC 1036
Db 639 ACCAGTGCAGGTCAGCAGCAGCAGCAGTGCATTCAGGAGCGCTGCTGTGTC 688
Qy 1037 GAGACTGCGCGCTGAGGAGCGCTGCGGAGCAGTTCGATGAGGAGGAGTGAAGT 1096
Db 689 GAGACTGCGCGCTGAGGAGCGCTGCGGAGCAGTTCGATGAGGAGGAGTGAAGT 748
Qy 1097 CCGTGGCTC 1105
Db 749 CCGTGGCTC 757
```

```
RESULT 6
BX332189 1070 bp mRNA linear EST 07-APR-2004
LOCUS BX332189 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
DEFINITION CDNA clone CS0DC008YL06 5-PRIME, mRNA sequence.
ACCESSION BX332189
VERSION BX332189.2 GI:46274586
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1070)
AUTHORS Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 2, 2003 this sequence version replaced gi:30341122.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
```

```
Email: seqref@genoscope.cns.fr Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6186.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?&CS0DC008DF03QPI&c=6186.f.
Location/Qualifiers
1..1070
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC008YL06"
/tissue="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

FEATURES

source

ORIGIN

```
Query Match 26.2%; Score 692; DB 4; Length 1070;
Best Local Similarity 99.6%; Pred. No. 6.9e-305;
Matches 842; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 838 CAACAGCAGAGCAGGAGCTGCAGAGTCTCTCGGGAGACAGAGCGTTCCTGAGCCAG 897
Db 1 CAACAGCAGAGCAGGAGCTGCAGAGTCTCTCGGGAGACAGAGCGTTCCTGAGCCAG 60
Qy 898 GTGCTAGCGCGGTGCTCAGCTGTCGCGCCAGCGGAGGTGCAGTCCACAAAGATGAG 957
Db 61 GTGCTAGCGCGGTGCTCAGCTGTCGCGCCAGCGGAGGTGCAGTCCACAAAGATGAG 120
Qy 958 GCCGTGTACCTGGGCCCTGAACAGCTGACGCTGAGCACCAGCAGCAAGTGCCTCATTTGCC 1017
Db 121 GCCGTGTACCTGGGCCCTGAACAGCTGACGCTGAGCACCAGCAGCAAGTGCCTCATTTGCC 180
Qy 1018 GAGCGCTGGTGTCTGTGCGAGACCTGCGCGCTGCGAGAGGCGCTCGGGAGACAGTGC 1077
Db 181 GAGCGCTGGTGTCTGTGCGAGACCTGCGCGCTGCGAGAGGCGCTCGGGAGACAGTGC 240
Qy 1078 ATGAGGAGGAGGAGTGAAGTGCCTGCTCAGCGCATCCCTCGCGGAGACATGCCCCACCA 1137
Db 241 ATGAGGAGGAGGAGTGAAGTGCCTGCTCAGCGCATCCCTCGCGGAGACATGCCCCACCA 300
Qy 1138 CTCATCCGCACCAACCGCTTCACGCGCCAGCTTCAGAGGCGATCGTGGATCGCTTACGCGCTG 1197
Db 301 CTCATCCGCACCAACCGCTTCACGCGCCAGCTTCAGAGGCGATCGTGGATCGCTTACGCGCTG 360
Qy 1198 GCGCGCTACAGAGAGGTCAACCGCGCTCCCTACACCATCATCATCTCCCTTCCTGTTT 1257
Db 361 GCGCGCTACAGAGAGGTCAACCGCGCTCCCTACACCATCATCATCTCCCTTCCTGTTT 420
Qy 1258 GCTGTGATGTTTCGGGATGTGCGGACAGCGGTGCTCATGTTCTCTCGCGCTGCGCATG 1317
Db 421 GCTGTGATGTTTCGGGATGTGCGGACAGCGGTGCTCATGTTCTCTCGCGCTGCGCATG 480
Qy 1318 GTCTTTGCGGAGAACCGACCGGCTGTGAAGCGCGCAGAGACGAGATCTGGCAGACTTTC 1377
Db 481 GTCTTTGCGGAGAACCGACCGGCTGTGAAGCGCGCAGAGACGAGATCTGGCAGACTTTC 540
Qy 1378 TTCAGGCGCGGTACCTGCTCTGCTTATGGGCGTGTTCCTCATCTACACCGGCTTCATC 1437
Db 541 TTCAGGCGCGGTACCTGCTCTGCTTATGGGCGTGTTCCTCATCTACACCGGCTTCATC 600
Qy 1438 TACAAAGAGTGTTCAGTGCAGCGCCACAGCATCTTCCTCGGCTGGGAGTGTGCGCGC 1497
Db 601 TACAAAGAGTGTTCAGTGCAGCGCCACAGCATCTTCCTCGGCTGGGAGTGTGCGCGC 660
Qy 1498 ATGCGCAACAGTGTGGCTGAGTGAATCATCTTCCTGGCGCCAGCACAGATGCTTACCTG 1557
```

```
Db 661 ATGCCAACCAAGTCTGGCTGGAGTGATGCTTCTTGGCCAGCACAGATCTTACCCCTG 720
QY 1558 GATCCCAACGTCACCGGTGCTTCTTGGGACCCCTACCCCTTTGGCATCGATCTTATTTGG 1617
Db 721 GATCCCAACGTCACCGGTGCTTCTTGGGACCCCTACCCCTTTGGCATCGATCTTATTTGG 780
QY 1618 AGCTGCTGCTCCACCACTTGGCTTCTTCAACTCTTCAAGTGAAGATGTCGGTCACTC 1677
Db 781 AGCTGCTGCTCCACCACTTGGCTTCTTCAACTCTTCAAGTGAAGATGTCGGTCACTC 840
QY 1678 CTGGG 1682
Db 841 CTGGG 845

RESULT 7
CR599870
LOCUS
DEFINITION
full-length cDNA clone CSODI006Y009 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION
CR599870
VERSION
HTC; CDSLT_cDNA.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 (bases 1 to 671)
AUTHORS
Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 671)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoK v sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1..671
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI006Y009"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 25.4%; Score 671; DB 6; Length 671;
Best Local Similarity 100.0%; Pred. No. 2.9e-295;
Matches 671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1953 GGCCAGCTGGTGGTCTGGCTTGGCCATGGTGGCCATCTCTGCTTGGGACACCCCT 2012
Db 1 GGCCAGCTGGTGGTCTGGCTTGGCCATGGTGGCCATCTCTGCTTGGGACACCCCT 60
QY 2013 GCACCTGCTGCACCGCCACCGCCCTGGCGAGGAGGCCCTGACCGACAGAGGA 2072
Db 61 GCACCTGCTGCACCGCCACCGCCCTGGCGAGGAGGCCCTGACCGACAGAGGA 120
QY 2073 AACCAAGGCGGGTGTCTGACCTGCTGACGATCTGTGAATGGTGGAGCTCCGATGA 2132
Db 121 AACCAAGGCGGGTGTCTGACCTGCTGACGATCTGTGAATGGTGGAGCTCCGATGA 180
QY 2133 GGNAAGGCGGGGGCTGGATGATGAGAGGAGCGGAGCTCGTCCCTCCGAGGTGCT 2192
```

```
Db 181 GGAAGAGGCGAGGGGCGCTGGATGATGAAGAGGAGGCGGAGCTCGTCCCTCCGAGGTGCT 240
QY 2193 CATGCACAGCGCATCCACACCATCGAGTCTTCTGCTGGGCTGCTCTCCACACCGCCTC 2252
Db 241 CATGCACAGCGCATCCACACCATCGAGTCTTCTGCTGGGCTGCTCTCCACACCGCCTC 300
QY 2253 CTACTCTGGCTGCTGGGCGCTGAGCTTGGCCACCGCCAGCTGCTCCGAGGTCTTGTGGC 2312
Db 301 CTACTCTGGCTGCTGGGCGCTGAGCTTGGCCACCGCCAGCTGCTCCGAGGTCTTGTGGC 360
QY 2313 CATGCTGATGGCATAGGCTTGGGCGCTGGGCGGAGGTGGGCTGGGCTGGGCTGCT 2372
Db 361 CATGCTGATGGCATAGGCTTGGGCGCTGGGCGGAGGTGGGCTGGGCTGCTGGTGTGCT 420
QY 2373 GGTCCCATCTTTGGCCGCTTGGCGGTGATGACCGTGGCTATCTCTGCTGGTGGAGGG 2432
Db 421 GGTCCCATCTTTGGCCGCTTGGCGGTGATGACCGTGGCTATCTCTGCTGGTGGAGGG 480
QY 2433 ACTCTAGCCTTCTGACCGCTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2492
Db 481 ACTCTAGCCTTCTGACCGCTTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 2493 CTCAGGACCGGCTACCAAGCTGAGTCCCTTCACTTCCCTGCTGCTGCTGCTGCTGCTGCT 2552
Db 541 CTCAGGACCGGCTACCAAGCTGAGTCCCTTCACTTCCCTGCTGCTGCTGCTGCTGCTGCT 600
QY 2553 ACTGCAAGTCTTCCGACAGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2612
Db 601 ACTGCAAGTCTTCCGACAGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 2613 CCGCCCTGGCA 2623
Db 661 CCGCCCTGGCA 671

RESULT 8
BQ073186
LOCUS
DEFINITION
BQ073186 1088 bp mRNA linear EST 02-APR-2002
AGENCOURT 6817733 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:575559
5', mRNA sequence.
ACCESSION
BQ073186
VERSION
BQ073186.1 GI:19902232
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE
1 (bases 1 to 1088)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM12799 row: p column: 08
High quality sequence stop: 593.
FEATURES
source
1..1088
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:575559"
/tissue_type="leukocyte"
/lab_host="DH10B"
/clone_lib="NIH_MGC_118"
```

/notes="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source leukocytes from anonymous pool of non-activated adult donors. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 1.2-3.3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 027. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 25.4%; Score 671; DB 3; Length 1088;
Best Local Similarity 100.0%; Pred. No. 2.8e-295;
Matches 671; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1919 ACAGGCTGCTCTACCCCGGAGGAGGTGTCAGGCCACGCTGTGTGCTTGGCTTGG 1978
DB 14 ACAGGCTGCTCTACCCCGGAGGAGGTGTCAGGCCACGCTGTGTGCTTGGCTTGG 73

QY 1979 CCATGCTGCCATCTCTGCTGTGGCACACCCCTGACCTGTGACCCGCCACCGCGCC 2038
DB 74 CCATGCTGCCATCTCTGCTGTGGCACACCCCTGACCTGTGACCCGCCACCGCGCC 133

QY 2039 GCTCGGAGGAGGCGCTGTGACCGACAGGAGAAACAAGCGCGGTGCTGGACCTGC 2098
DB 134 GCTCGGAGGAGGCGCTGTGACCGACAGGAGAAACAAGCGCGGTGCTGGACCTGC 193

QY 2099 CTGACGCATCTGTGAATGCTGTGAGTCCGATGAGAAAGAGGAGGCGCTGTGATG 2158
DB 194 CTGACGCATCTGTGAATGCTGTGAGTCCGATGAGAAAGAGGAGGCGCTGTGATG 253

QY 2159 AAGAGGAGGCGAGCTGCTCCCTCGAGGTGCTCATGACACAGGCGCATCCACCATCG 2218
DB 254 AAGAGGAGGCGAGCTGCTCCCTCGAGGTGCTCATGACACAGGCGCATCCACCATCG 313

QY 2219 AGTTCTGCTGGGCTGCTCTCAACACCGCTCTTACCTGCGCCTGTGGCGCTGAGCC 2278
DB 314 AGTTCTGCTGGGCTGCTCTCAACACCGCTCTTACCTGCGCCTGTGGCGCTGAGCC 373

QY 2279 TGGCCACGCCAGCTGCTCGAGTCTGTGGCCATGTGTGATGCGCATAGGCTGGGCC 2338
DB 374 TGGCCACGCCAGCTGCTCGAGTCTGTGGCCATGTGTGATGCGCATAGGCTGGGCC 433

QY 2339 TGGGCGGAGGTGGGCGTGGCGCTGTGTGCTGCTCCCATCTTTGCGCGCTTGGCG 2398
DB 434 TGGGCGGAGGTGGGCGTGGGCGTGTGTGCTGCTGCTCCCATCTTTGCGCGCTTGGCG 493

QY 2399 TGATGACCGTGTCTATCTGCTGTGTGATGAGGAGCTCTCAGCTTCTTGACGCGCTGC 2458
DB 494 TGATGACCGTGTCTATCTGCTGTGTGATGAGGAGCTCTCAGCTTCTTGACGCGCTGC 553

QY 2459 GGCTGCACTGGTGGATTCAGACACAGTCTTACTCAGGACGCGGCTACAGCTGATC 2518
DB 554 GGCTGCACTGGTGGATTCAGACACAGTCTTACTCAGGACGCGGCTACAGCTGATC 613

QY 2519 CCTTACCTTCGTGCGCAGATGACTAGGCGGCTCTGAGGTCTGCGCAGACCTCTTC 2578
DB 614 CCTTACCTTCGTGCGCAGATGACTAGGCGGCTCTGAGGTCTGCGCAGACCTCTTC 673

QY 2579 CTGACCTCTGA 2589
DB 674 CTGACCTCTGA 684

RESULT 9

BU732254/c
LOCUS BU732254
DEFINITION BU732254 664 bp mRNA linear EST 09-OCT-2002
UI-E-COI-afx-k-03-0-UI.81 UI-E-COI Homo sapiens cdna clone
ACCESSION BU732254
VERSION BU732254
KEYWORDS BU732254.1 GI:23657965
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 664)

Bonaldo, M.P., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

8889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue procurement: Dr. Gregg Hageman

cdna Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES

Location/Qualifiers

1..664

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-E-COI-afx-k-03-0-UI"

/tissue_type="optic nerve"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-E-COI"

/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

UI-E-COI is a normalized cDNA library containing the

following tissue(s): optic nerve. The library was

constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA

synthesis was primed with an oligo-dT primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I

adaptor, digested with Not I, and cloned directionally

into pT73-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag

sequence that is located between the Not I site and the

(dT)18 tail. The sequence tag for this library is

CCATTAAGT. This library was created for the program, Gene

discovery in the Visual System, supported by National Eye

Institute (NEI).

TAG_TISSUE=human optic nerve

TAG_LIB=UI-E-COI

TAG_SEQ=CCATTAAGT

ORIGIN

Query Match 23.5%; Score 620; DB 3; Length 664;
Best Local Similarity 100.0%; Pred. No. 6.2e-272;
Matches 620; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2021 TGCACCGCCACCGCCCGCTGCGAGGAGGAGCCCGCTGACCGACAGGAGGAAACAAG 2080
DB 620 TGCACCGCCACCGCCCGCTGCGAGGAGGAGCCCGCTGACCGACAGGAGGAAACAAG 561

QY 2081 CCGGGTGTGTGACCTGTGCTGACGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGG 2140
DB 560 CCGGGTGTGTGACCTGTGCTGACGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGG 501

QY 2141 CAGGGGGCTGTGATGATGAAGAGGAGGCGGAGCTGTCCCTCCGAGGTGCTCATGCACC 2200
DB 500 CAGGGGGCTGTGATGATGAAGAGGAGGCGGAGCTGTCCCTCCGAGGTGCTCATGCACC 441

QY 2201 AGGCCATCCACACCATCGAGTTCTGCTGGGCTGTCTCCACACCGCCTCTCTACTCTG 2260

```
Db 440 AGGCATCCACACCATCGAGTTCTGGCTGGGCTGGTCTCAACACCGCTCTTACCTGC 381
Qy 2261 GCTGTGGGCTTGGAGCTGGCCACGCCAGCTGTCCGAGGTTCTGTGGGCCATGTGTA 2320
Db 380 GCCTGTGGGCTTGGAGCTGGCCACGCCAGCTGTCCGAGGTTCTGTGGGCCATGTGTA 321
Qy 2321 TGGGCATAGGCTGGGCTGGGCGGAGGAGTGGGCGTGGGCTGTGGTGTGTTGTCCTCCCA 2380
Db 320 TGGGCATAGGCTGGGCTGGGCGGAGGAGTGGGCGTGGGCTGTGGTGTGTTGTTGTCCTCCCA 261
Qy 2381 TCTTTGGCGCTTTGGCGTGTGATGACCTGGTGTATCTCTGCTGGTGTGATGAGGAGTCTCAG 2440
Db 260 TCTTTGGCGCTTTGGCGTGTGATGACCTGGTGTATCTCTGCTGGTGTGATGAGGAGTCTCAG 201
Qy 2441 CTTTCTGTGACGCCCTGGCTGGCTGACCTGGTGTGATGATTCAGAACAACTTCTACTCAGGCA 2500
Db 200 CTTTCTGTGACGCCCTGGCTGGCTGACCTGGTGTGATGATTCAGAACAACTTCTACTCAGGCA 141
Qy 2501 CGGGCTACAGCTGAGTCCCTTACCTTCCCTGCTGCTGACAGATGATAGGCGCCCTGAGG 2560
Db 140 CGGGCTACAGCTGAGTCCCTTACCTTCCCTGCTGCTGACAGATGATAGGCGCCCTGAGG 81
Qy 2561 TCTTGCAGACCTCTCTCTGACCTCTGAGGCGAGGAGGATTAAGACAGGTCGCGCTG 2620
Db 80 TCTTGCAGACCTCTCTCTGACCTCTGAGGCGAGGAGGATTAAGACAGGTCGCGCTG 21
Qy 2621 GCACCAAAAAAAAAAAAA 2640
Db 20 GCACCAAAAAAAAAAAAA 1
```

RESULT 10

CA425567/c

LOCUS

DEFINITION

UI-H-DF0-bem-m-06-0-UI.s1 NCI CGAP DF0 Homo sapiens cDNA clone

UI-H-DF0-bem-m-06-0-UI 3', mRNA sequence.

CA425567

CA425567.1 GI:24788293

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 619)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaps-x@mail.nih.gov

Tissue Procurement: Dr. Jose Mercuende

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, bento-soares@uiowa.edu

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers

1. .619

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-H-DF0-bem-m-06-0-UI"

/tissue_type="Subchondral Bone"

/dev_stage="Adult"

/lab_host="PHI08 (Life Technologies)"

/clone_lib="NCI CGAP DF0"

/notes="Organ: Bone; Vector: pT73-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

NCI-CGAP_DF0 is a cDNA library containing the following

tissue(s): Subchondral Bone. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor. Digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GTTAAAGCGTC.

TAG_TISSUE=subchondral bone

TAG_LIB=UI-H-DF0

TAG_SEQ=GTAAAGCGTC"

ORIGIN

```
Query Match 23.4%; Score 619; DB 4; Length 619;
Best Local Similarity 100.0%; Pred. No. 1.8e-271;
Matches 619; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2022 GCACCGGCACCGCGCGCTGCGGAGGAGGCGCGCTGACGACAGAGGAGGAGGAGGAGG 2081
Db 619 GCACCGGCACCGCGCGCTGCGGAGGAGGCGCGCTGACGACAGAGGAGGAGGAGGAGG 560
Qy 2082 CGGGTTGCTGGACCTGCCTGACGATCTGTGAATGGCTGGAGCTCCGATGAGGAGGAGG 2141
Db 559 CGGGTTGCTGGACCTGCCTGACGATCTGTGAATGGCTGGAGCTCCGATGAGGAGGAGG 500
Qy 2142 AGGGGGCTGCGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2201
Db 499 AGGGGGCTGCGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 440
Qy 2202 GGCATCCACACCATCGAGTTCTGCTGGGCTGGCTTCCAAACACCGCTCTTACCTGGC 2261
Db 439 GGCATCCACACCATCGAGTTCTGCTGGGCTGGCTTCCAAACACCGCTCTTACCTGGC 380
Qy 2262 CTTGTGGGCTTGGAGCTGGCCACCGCCAGCTGCGAGGTTCTGCGAGGTTCTGCGGCTG 2321
Db 379 CTTGTGGGCTTGGAGCTGGCCACCGCCAGCTGCGAGGTTCTGCGAGGTTCTGCGGCTG 320
Qy 2322 GCGCATAGGCTGGGCTGGGCGGGAGGTTGGGCGGCTGTGCTGGTGGTGGTGGTGGTGG 2381
Db 319 GCGCATAGGCTGGGCTGGGCGGGAGGTTGGGCGGCTGTGCTGGTGGTGGTGGTGGTGG 260
Qy 2382 CTTTGGCGCTTTGGCGTGATGACCGTGGCTATCTCTGCTGGTGGTGGTGGTGGTGGTGG 2441
Db 259 CTTTGGCGCTTTGGCGTGATGACCGTGGCTATCTCTGCTGGTGGTGGTGGTGGTGGTGG 200
Qy 2442 CTTTCTGACGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 2501
Db 199 CTTTCTGACGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 140
Qy 2502 GGGCTACAGCTGAGTCCCTTACCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 2561
Db 139 GGGCTACAGCTGAGTCCCTTACCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 80
Qy 2562 CTTGCGCAGCTCTCTTCTGACCTCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2621
Db 79 CTTGCGCAGCTCTCTTCTGACCTCTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 20
Qy 2622 CAAAAA 2640
Db 19 CAAAAA 1
```

RESULT 11

CR992406

LOCUS

DEFINITION

CR992406 RZPD no.9016 Homo sapiens cDNA clone RZPDp901611642 5',

mRNA sequence.

ACCESSION

CR992406

VERSION

CR992406.1 GI:68286291

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

CR992406 842 bp mRNA linear EST 28-JUN-2005
CR992406 RZPD no.9016 Homo sapiens cDNA clone RZPDp901611642 5',
mRNA sequence.

ACCESSION CR992406
VERSION CR992406.1
KEYWORDS GI:68286291
EST. EST.
SOURCE Homo sapiens (human)

ORIGIN Intramural Sequencing Center (NISC)."

Query Match 22.8%; Score 603; DB 5; Length 603;
Best Local Similarity 100.0%; Pred. No. 3.8e-264;
Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 574 GCCGTGAGCCCAACAGGCCCCCTGAGAGCGCTGCTCTGAGGGCCCTGCGGGC 633
DB 1 GCCGTGAGCCCAACAGGCCCCCTGAGAGCGCTGCTCTGAGGGCCCTGCGGGC 60

QY 634 TTCCTCATATGACGCTTTCAGGAGCTGAGCAGCGCTGAGCAGCCCGTGCAGCGGCAG 693
DB 61 TTCCTCATATGACGCTTTCAGGAGCTGAGCAGCGCTGAGCAGCCCGTGCAGCGGCAG 120

QY 694 CGAGCCAGTGGATGACCTTCTCATCTCTCTTCTGAGGAGCAGATCGGACAGATC 753
DB 121 CCAGCCAGTGGATGACCTTCTCATCTCTCTTCTGAGGAGCAGATCGGACAGATC 180

QY 754 CGCAAGATCAGGAGCTGCTTCCAGTCCAGCTCTTCCGCTTCTGAGCAGCGAGGAGGCC 813
DB 181 CGCAAGATCAGGAGCTGCTTCCAGTCCAGCTCTTCCGCTTCTGAGCAGCGAGGAGGCC 240

QY 814 CGCTCGGGGCCCTGACAGAGCTGCAACAGCAGAGCCAGAGCTGCAAGGAGTCTCGGG 873
DB 241 CGCTCGGGGCCCTGACAGAGCTGCAACAGCAGAGCCAGAGCTGCAAGGAGTCTCGGG 300

QY 874 GAGACAGCGGTTCTGAGCAGGTTCTAGGCGGCTGCTGAGCTGCTGCTGCGCCAGGG 933
DB 301 GAGACAGCGGTTCTGAGCAGGTTCTAGGCGGCTGCTGAGCTGCTGCTGCGCCAGGG 360

QY 934 CAGGTGAGCTCCACAGATGAGGCGCTGCTGAGCTGCTGAGCAGCTGAGCGTGGC 993
DB 361 CAGGTGAGCTCCACAGATGAGGCGCTGCTGAGCTGCTGAGCAGCTGAGCGTGGC 420

QY 994 ACCACGACAAAGTGCCTCATTTGCGAGGCTGCTGCTGCGAGACCTGCGCGCCCTG 1053
DB 421 ACCACGACAAAGTGCCTCATTTGCGAGGCTGCTGCTGCGAGACCTGCGCGCCCTG 480

QY 1054 CAGGAGCCCTGCGGAGCAGCTGATGAGGAGGAGTGTAGTGGCTTACCGCATC 1113
DB 481 CAGGAGCCCTGCGGAGCAGCTGATGAGGAGGAGTGTAGTGGCTTACCGCATC 540

QY 1114 CCTGCGGAGACATGCCCCCAGCTCATCTCGACACACCGCTTACGCGCAGCTTCAG 1173
DB 541 CCTGCGGAGACATGCCCCCAGCTCATCTCGACACACCGCTTACGCGCAGCTTCAG 600

QY 1174 GGC 1176
DB 601 GGC 603

RESULT 13
CR986883
LOCUS
DEFINITION CR986883 RZPD no.9016 Homo sapiens cdna clone RZPDp9016E0116 5',
mRNA sequence.
ACCESSION CR986883
VERSION CR986883.1 GI:68280768
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 699)
AUTHORS Heil, O., Ebert, L., Hennig, S., Henze, S., Radelof, U., Schneider, D.
and Korn, B.
TITLE Human T-Lymphocytes library
JOURNAL Unpublished (2005)
COMMENT RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany

Email: www.rzpd.de
RZPD; RZPDp9016E0116.
RZPDlib; (Human T-Lymphocytes) RZPD LIB No.9016
http://www.rzpd.de/cgi-bin/products/set.cgi?libNo=9016 Contact:
Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=RZPDp9016E0116
contact RZPD (product- support@rzpd.de) for further information.
Primer name: ges_4 , Primer sequence: CGATAACAATTCACACAG.

FEATURES

source

1. .699
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="RZPDp9016E0116"
/tissue_type="T-Lymphocytes"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RZPD no.9016"
/note="Vector: pQE80LSN_cloned; Site_1: Salt; Site_2:
NotI; vector:
http://www.rzpd.de/info/vectors/pQE80LSN_cloned_pic.shtml
; 1st strand cDNA was prepared from mRNA obtained from
human T-Lymphocytes with a NotI - oligo(dT) primer [5'
GACTAGTCTAGATCGAGCGCGCTTTTCTTTTCTTTT 3']
Double-stranded cDNA was ligated to SalI adaptors,
digested with NotI and cloned into the NotI and SalI sites
of the pQE80LSN_cloned vector"

ORIGIN

Query Match 22.8%; Score 602; DB 8; Length 699;
Best Local Similarity 100.0%; Pred. No. 1.1e-263;
Matches 602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 561 CTTTGTGGCAGGTGCGGTGGAGCCCAACAGGCCCCCTGAGCGCCTGCTGCTGGAG 620
DB 1 CTTTGTGGCAGGTGCGGTGGAGCCCAACAGGCCCCCTGAGCGCCTGCTGCTGGAG 60

QY 621 GGCTGCGCGGCTTCTCTATTCGACAGCTTCAGGAGCTGGAGCGCGCTGGAGCACC 680
DB 61 GGCTGCGCGGCTTCTCTATTCGACAGCTTCAGGAGCTGGAGCGCGCTGGAGCACC 120

QY 681 CGTGACGGCGAGCAGCAGCGTGGATGACCTTCTCATCTCTTACTGGGTGAGCAGAT 740
DB 121 CGTGACGGCGAGCAGCAGCGTGGATGACCTTCTCATCTCTTACTGGGTGAGCAGAT 180

QY 741 CGGACAGAGATCCGCAAGATCAGGAGCTGCTTCCACTGCGCAGCTTTCCTGCTTCTGCA 800
DB 181 CGGACAGAGATCCGCAAGATCAGGAGCTGCTTCCACTGCGCAGCTTTCCTGCTTCTGCA 240

QY 801 GCAGGAGAGGCGCGCTCGGGCCCTGCGAGAGCTGCAACAGCAGAGCAGAGCTGCA 860
DB 241 GCAGGAGAGGCGCGCTCGGGCCCTGCGAGAGCTGCAACAGCAGAGCAGAGCTGCA 300

QY 861 GGAGGTCTCGGGGAGACAGAGCGGTTCCTGAGCCAGGTGCTAGGCGGGTGTGCACT 920
DB 301 GGAGGTCTCGGGGAGACAGAGCGGTTCCTGAGCCAGGTGCTAGGCGGGTGTGCACT 360

QY 921 GCTGCGCCAGGCGCAGGTGCGAGTCCACAGAGTGAAGCGCGTGTACTGCGCCCTGAACCA 980
DB 361 GCTGCGCCAGGCGCAGGTGCGAGTCCACAGAGTGAAGCGCGTGTACTGCGCCCTGAACCA 420

QY 981 GTGACAGGTGAGCACAACAGCAAGTGCCTCATTTGCGAGGCTGTGTGTCGAGA 1040
DB 421 GTGACAGGTGAGCACAACAGCAAGTGCCTCATTTGCGAGGCTGTGTGTCGAGA 480

QY 1041 CCTGCGCGCCCTGAGAGGCGCCCTGCGGAGCAGCTCGATGAGGAGGAGTGTGTCGCT 1100

Db 481 CCTGCCCGCCCTGCGAGGCGCCCTGCGGACAGCTCGATGAGGAGGAGTGAGTGCCTG 540
 QY 1101 GGTACCCGATCCCTGCGGGACATGCCCCCCCCACACTCCGACCAACCGCTTAC 1160
 Db 541 GGTACCCGATCCCTGCGGGACATGCCCCCCCCACACTCATCCGACCAACCGCTTAC 600
 QY 1161 GG 1162
 Db 601 GG 602

RESULT 14
 BU630675/c
 LOCUS
 DEFINITION
 UI-H-FLO-bdm-n-01-0-UI.s1 NCI_CGAP_FLO Homo sapiens cDNA clone
 UI-H-FLO-bdm-n-01-0-UI 3', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue procurement: James Martin

CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 Seq primer: M13 FORWARD
 POLYA=Yes

FEATURES
 source
 1..618
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-H-FLO-bdm-n-01-0-UI"
 /tissue_type="Cell lines"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP_FLO"
 /note="Organ: Chondrosarcoma; Vector: pT7T3-Pac
 (Pharmacia) with a modified polylinker; Site_1: Ecor I;
 Site_2: Not I; NCI_CGAP_FLO is a cDNA library derived from
 a pool of mRNA obtained from 4 cell lines from grade III
 chondrosarcoma tissues. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an Ecor I adaptor, digested
 with Not I, and cloned directionally into pT7T3-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is GAGGTCGGT. The cell line
 was provided by Dr. James Martin from University of Iowa.
 TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
 TAG_LIB=UI-H-FLO
 TAG_SEQ=GAGGTCGGTG"

ORIGIN

Query Match 22.7%; Score 599; DB 3; Length 618;
 Best Local Similarity 100.0%; Pred. No. 2.5e-262;
 Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2042 TCGCAGGAGGCGCCCTGACGACAGGAGGAAAAAAGAGCCGGTTGTGGACCTGCCTG 2101
 Db 599 TCGCAGGAGGCGCCCTGACGACAGGAGGAAAAAAGAGCCGGTTGTGGACCTGCCTG 540
 QY 2102 AGCGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAAAGGAGCGGCTTGGATGATGAAG 2161
 Db 539 AGCGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAAAGGAGCGGCTTGGATGATGAAG 480
 QY 2162 AGGAGGCGAGCTCGTCCCTCCGAGGTGCTCATGACACAGGCCATTCACACATCGAGT 2221
 Db 479 AGGAGGCGAGCTCGTCCCTCCGAGGTGCTCATGACACAGGCCATTCACACATCGAGT 420
 QY 2222 TCTGCTGGGCTGCGTCTCCAAACCGCTCTACTCGCTGCGCTTGGGCGCTGAGCCTG 2281
 Db 419 TCTGCTGGGCTGCGTCTCCAAACCGCTCTACTCGCTGCGCTTGGGCGCTGAGCCTG 360
 QY 2282 CCCACGCCAGCTGTCCGAGGTTCCTGTGGGCCATGATGCGCATAGGCCCTGGGCGCTG 2341
 Db 359 CCCACGCCAGCTGTCCGAGGTTCCTGTGGGCCATGATGCGCATAGGCCCTGGGCGCTG 300
 QY 2342 GCGCGGAGGTGGGCGGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 2401
 Db 299 GCGCGGAGGTGGGCGGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 240
 QY 2402 TGACCGTGGCTATCTGCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 2461
 Db 239 TGACCGTGGCTATCTGCTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 180
 QY 2462 TGACCTGGGTGGAATTCAGAACAAAGTTCTACTCAGGACGCGGCTACAGCTGAGTCCCT 2521
 Db 179 TGACCTGGGTGGAATTCAGAACAAAGTTCTACTCAGGACGCGGCTACAGCTGAGTCCCT 120
 QY 2522 TCACCTTGTCTGCCACAGATGATAGGCGCCACTGAGGTCTCTGCCAGACCTCTCTCTCTG 2581
 Db 119 TCACCTTGTCTGCCACAGATGATAGGCGCCACTGAGGTCTCTGCCAGACCTCTCTCTCTG 60
 QY 2582 ACCTCTGAGGCGAGGAGGAAATAAGAGCGTCCGCTGGCAAAAAAAGAAAAA 2640
 Db 59 ACCTCTGAGGCGAGGAGGAAATAAGAGCGTCCGCTGGCAAAAAAAGAAAAA 1

RESULT 15
 BU631560/c
 LOCUS
 DEFINITION
 UI-H-FLO-bdo-e-04-0-UI.s1 NCI_CGAP_FLO Homo sapiens cDNA clone
 UI-H-FLO-bdo-e-04-0-UI 3', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homiidae; Homo.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue procurement: James Martin

CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 Seq primer: M13 FORWARD
 POLYA=Yes

FEATURES
 source
 1..619
 /organism="Homo sapiens"
 /mol_type="mRNA"

```
/db_xref="taxon:9606"  
/clone="UI-H-FL0-bdo-e-04-0-UI"  
/tissue_type="Cell lines"  
/dev_stage="Adult"  
/lab_host="DH10B (Life Technologies)"  
/clone_lib="NCI CGAP FL0"  
/note="Organ: Chondrosarcoma; Vector: p77T3-Pac  
(Pharmacia) with a modified polylinker; Site 1: Ecor I;  
Site 2: Not I; NCI CGAP FL0 is a cDNA library derived from  
a pool of mRNA obtained from 4 cell lines from grade III  
chondrosarcoma tissues. The library was constructed  
according to Bonaldo, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an Ecor I adaptor, digested  
with Not I, and cloned directionally into p77T3-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tag for this library is GAGTTCGGTG. The cell line  
was provided by Dr James Martin from University of Iowa.  
TAG TISSUE=Human Chondrosarcoma Grade 3 cell line mix  
TAG LIB=UI-H-FL0  
TAG_SEQ=GAGTTCGGTG"
```

ORIGIN

```
Query Match      22.7%; Score 599; DB 3; Length 619;  
Best Local Similarity 100.0%; Pred. No. 2.5e-262;  
Matches 599; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2042 TCGGAGGAGCCGCTGACCGAGGAGGAAACAGCCGGGTGCTGGACCTGGCTG 2101  
DB      |||||||  
QY 599  TCGGAGGAGCCGCTGACCGAGGAGGAAACAGCCGGGTGCTGGACCTGGCTG 540  
DB      |||||||  
QY 2102 ACGCATCTGTAATGGCTGGAGCTCCGATGAGGAAAGGACAGGGGGCTGGATGATGAAG 2161  
DB      |||||||  
QY 539  ACGCATCTGTAATGGCTGGAGCTCCGATGAGGAAAGGACAGGGGGCTGGATGATGAAG 480  
DB      |||||||  
QY 2162 AGGAGGCGGAGCTGCTCCCTCCGAGGTCTCATGACCGAGGCTCCACACCATCGAGT 2221  
DB      |||||||  
QY 479  AGGAGGCGGAGCTGCTCCCTCCGAGGTCTCATGACCGAGGCTCCACACCATCGAGT 420  
DB      |||||||  
QY 2222 TCTGCTCTGGGCTGGCTCCAAACACCGCTCTACCTGGCGCTGGGGCTGGAGCTGG 2281  
DB      |||||||  
QY 419  TCTGCTCTGGGCTGGCTCTCAACACCGCTCTACCTGGCGCTGGGGCTGGAGCTGG 360  
DB      |||||||  
QY 2282 CCCACGCCAGCTCTCCGAGGTTCTGTGGGCCATGTGTGATGCGCATAGGCTGGGCTGG 2341  
DB      |||||||  
QY 359  CCCACGCCAGCTCTCCGAGGTTCTGTGGGCCATGTGTGATGCGCATAGGCTGGGCTGG 300  
DB      |||||||  
QY 2342 GCGGGGAGGTGGGGCTGGGGCTGTGTGTGTCCTCCCATCTTTGGCGCTTTGGCGTGA 2401  
DB      |||||||  
QY 299  GCGGGGAGGTGGGGCTGGGGCTGTGTGTGTCCTCCCATCTTTGGCGCTTTGGCGTGA 240  
DB      |||||||  
QY 2402 TGACCGTGGCTATCTCTGCTGGTATGAGGAGCTCTCAGCTTCCTGCACGCCCTGGCGC 2461  
DB      |||||||  
QY 239  TGACCGTGGCTATCTCTGCTGGTATGAGGAGCTCTCAGCTTCCTGCACGCCCTGGCGC 180  
DB      |||||||  
QY 2462 TGCACTGGGTGGAAATTCAGAAACAAGTTCTACTCAGGCACGGGCTACAAGCTGAGTCCCT 2521  
DB      |||||||  
QY 179  TGCACTGGGTGGAAATTCAGAAACAAGTTCTACTCAGGCACGGGCTACAAGCTGAGTCCCT 120  
DB      |||||||  
QY 2522 TCACCTTCGGCTGCCACAGATGACTAGGGGCCACCTGCAGGTCTCTGCCAGACTCTCTCTG 2581  
DB      |||||||  
QY 119  TCACCTTCGGCTGCCACAGATGACTAGGGGCCACCTGCAGGTCTCTGCCAGACTCTCTCTG 60  
DB      |||||||  
QY 2582 ACCTCTGAGCGAGGAGGAATTAAGACGCTCCGCCCTGGCAAAAAAAAAAAAAA 2640  
DB      |||||||  
QY 59  ACCTCTGAGCGAGGAGGAATTAAGACGCTCCGCCCTGGCAAAAAAAAAAAAAA 1
```

Search completed: June 30, 2006, 09:14:19
Job time : 12287 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2006, 05:49:39 ; Search time 490 Seconds
(without alignments)
10081.076 Million cell updates/sec

Title: US-10-783-519-1

Perfect score: 2640

Sequence: 1 cggcgctgcggacggcgag.....gcaaaaaaaaaaaaaaaaaa 2640

Scoring table: OLOG NUC

Gapop 60.0 , Gapext 60.0

Searched: 1403666 seqs, 935554401 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2806514

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /EMC_Celerra_SID33/ptodata/2/ina/1_COMB.seq.*
- 2: /EMC_Celerra_SID33/ptodata/2/ina/5_COMB.seq.*
- 3: /EMC_Celerra_SID33/ptodata/2/ina/6A_COMB.seq.*
- 4: /EMC_Celerra_SID33/ptodata/2/ina/6B_COMB.seq.*
- 5: /EMC_Celerra_SID33/ptodata/2/ina/7_COMB.seq.*
- 6: /EMC_Celerra_SID33/ptodata/2/ina/H_COMB.seq.*
- 7: /EMC_Celerra_SID33/ptodata/2/ina/PCUS_COMB.seq.*
- 8: /EMC_Celerra_SID33/ptodata/2/ina/PP_COMB.seq.*
- 9: /EMC_Celerra_SID33/ptodata/2/ina/RE_COMB.seq.*
- 10: /EMC_Celerra_SID33/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 2640 | 100.0 | 2640 | 3 | US-08-684-932A-37 |
| 2 | 2640 | 100.0 | 2640 | 3 | US-09-618-304B-1 |
| 3 | 2640 | 100.0 | 2655 | 3 | US-09-016-434-1094 |
| 4 | 2640 | 100.0 | 2655 | 3 | US-09-023-655-916 |
| 5 | 2640 | 100.0 | 2655 | 4 | US-09-880-107-3363 |
| 6 | 2142 | 81.1 | 2654 | 3 | US-09-949-016-2404 |
| 7 | 2142 | 81.1 | 2654 | 3 | US-09-949-016-2405 |
| 8 | 2127 | 80.6 | 2706 | 3 | US-09-949-016-2975 |
| 9 | 2127 | 80.6 | 2706 | 3 | US-09-949-016-2976 |
| 10 | 1670 | 63.3 | 2457 | 3 | US-09-949-016-1330 |
| 11 | 1670 | 63.3 | 2457 | 3 | US-09-949-016-1331 |
| 12 | 442 | 16.7 | 493 | 4 | US-09-880-107-1291 |
| 13 | 221 | 8.4 | 601 | 3 | US-09-949-016-83496 |
| 14 | 221 | 8.4 | 601 | 3 | US-09-949-016-83505 |
| 15 | 221 | 8.4 | 601 | 3 | US-09-949-016-107576 |
| 16 | 221 | 8.4 | 601 | 3 | US-09-949-016-107585 |
| 17 | 221 | 8.4 | 11917 | 3 | US-09-949-016-13072 |
| 18 | 221 | 8.4 | 11917 | 3 | US-09-949-016-13073 |
| 19 | 221 | 8.4 | 15061 | 3 | US-09-949-016-14717 |
| 20 | 221 | 8.4 | 15061 | 3 | US-09-949-016-14718 |
| 21 | 221 | 8.4 | 15859 | 3 | US-09-949-016-14146 |
| 22 | 221 | 8.4 | 15859 | 3 | US-09-949-016-14147 |
| 23 | 160 | 6.1 | 601 | 3 | US-09-949-016-46215 |

24 160 6.1 601 3 US-09-949-016-46216 Sequence 46216, A
25 160 6.1 601 3 US-09-949-016-46222 Sequence 46222, A
26 160 6.1 601 3 US-09-949-016-46223 Sequence 46223, A
27 160 6.1 601 3 US-09-949-016-83501 Sequence 83501, A
28 160 6.1 601 3 US-09-949-016-83502 Sequence 83502, A
29 160 6.1 601 3 US-09-949-016-83510 Sequence 83510, A
30 160 6.1 601 3 US-09-949-016-83511 Sequence 83511, A
31 160 6.1 601 3 US-09-949-016-107581 Sequence 107581, A
32 160 6.1 601 3 US-09-949-016-107582 Sequence 107582, A
33 160 6.1 601 3 US-09-949-016-107590 Sequence 107590, A
34 160 6.1 601 3 US-09-949-016-107591 Sequence 107591, A
35 124 4.7 601 3 US-09-949-016-46214 Sequence 46214, A
36 124 4.7 601 3 US-09-949-016-46221 Sequence 46221, A
37 124 4.7 601 3 US-09-949-016-83500 Sequence 83500, A
38 124 4.7 601 3 US-09-949-016-83509 Sequence 83509, A
39 124 4.7 601 3 US-09-949-016-107580 Sequence 107580, A
40 124 4.7 601 3 US-09-949-016-107589 Sequence 107589, A
41 122 4.6 601 3 US-09-949-016-83495 Sequence 83495, A
42 122 4.6 601 3 US-09-949-016-83504 Sequence 83504, A
43 122 4.6 601 3 US-09-949-016-107575 Sequence 107575, A
44 122 4.6 601 3 US-09-949-016-107584 Sequence 107584, A
45 112 4.2 601 3 US-09-949-016-46217 Sequence 46217, A

ALIGNMENTS

RESULT 1
US-08-684-932A-37
; Sequence 37, Application US/08684932A
; Patent No. 6403304
; GENERAL INFORMATION:
; APPLICANT: Stashenko, Philip
; APPLICANT: Li, Yi-ping
; APPLICANT: Wuchertfennig, Anne L.
; TITLE OF INVENTION: HUMAN OSTEOCLAST-SPECIFIC AND -RELATED
; TITLE OF INVENTION: DNA SEQUENCES
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/684,932A
; FILING DATE: 19-JUL-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: FDC92-02FM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2640 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 58..2523
US-08-684-932A-37

| | | | | | | | | | |
|-----------------------|------|--------------|--------------|------------|--------------|----------|----------|----------|----------|
| Query Match | | 100.0%; | Score 2640; | DB 3; | Length 2640; | | | | |
| Best Local Similarity | | 100.0%; | Pred. No. 0; | | | | | | |
| Matches 2640; | | Conservative | 0; | Mismatches | 0; | Indels | 0; | Gaps | 0; |
| Qy | 1 | CGGCGTGC | CGGAGCGG | CGGAGCGG | CGGAGCGG | CGGAGCGG | CGGAGCGG | CGGAGCGG | CGGAGCGG |
| Db | 1 | CGGCGTGC | CGGAGCGG | CGGAGCGG | CGGAGCGG | CGGAGCGG | CGGAGCGG | CGGAGCGG | CGGAGCGG |
| Qy | 61 | GGCTCCAT | GGTCCG | GGAGGAGG | GGGCGG | GGGCGG | GGGCGG | GGGCGG | GGGCGG |
| Db | 61 | GGCTCCAT | GGTCCG | GGAGGAGG | GGGCGG | GGGCGG | GGGCGG | GGGCGG | GGGCGG |
| Qy | 121 | GCCTACAC | CTGCGT | GGGCGG | GGGCGG | GGGCGG | GGGCGG | GGGCGG | GGGCGG |
| Db | 121 | GCCTACAC | CTGCGT | GGGCGG | GGGCGG | GGGCGG | GGGCGG | GGGCGG | GGGCGG |
| Qy | 181 | GGCTCGGT | GAGCGG | CTTCCG | AGAGAG | AGAGAG | AGAGAG | AGAGAG | AGAGAG |
| Db | 181 | GGCTCGGT | GAGCGG | CTTCCG | AGAGAG | AGAGAG | AGAGAG | AGAGAG | AGAGAG |
| Qy | 241 | GAGAGAC | CTTCCG | GGGCGG | GGGCGG | GGGCGG | GGGCGG | GGGCGG | GGGCGG |
| Db | 241 | GAGAGAC | CTTCCG | GGGCGG | GGGCGG | GGGCGG | GGGCGG | GGGCGG | GGGCGG |
| Qy | 301 | CCAAAGGG | AGGCGT | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Db | 301 | CCAAAGGG | AGGCGT | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Qy | 361 | GAGCGG | CTGCGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Db | 361 | GAGCGG | CTGCGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Qy | 421 | CTGACAC | AGTGGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Db | 421 | CTGACAC | AGTGGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Qy | 481 | GCGCGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Db | 481 | GCGCGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Qy | 541 | CACGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Db | 541 | CACGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Qy | 601 | CTAGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Db | 601 | CTAGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Qy | 661 | GAGCAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Db | 661 | GAGCAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Qy | 721 | TGCTACT | GGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Db | 721 | TGCTACT | GGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Qy | 781 | CAGCTCT | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Db | 781 | CAGCTCT | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Qy | 841 | CAGCAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Db | 841 | CAGCAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Qy | 901 | CTAGGCG | GGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Db | 901 | CTAGGCG | GGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Qy | 961 | GTGTACT | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Db | 961 | GTGTACT | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Qy | 1021 | GGCTGGT | GCTCTG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |
| Db | 1021 | GGCTGGT | GCTCTG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG | CGGAGG |

| | | | |
|----|------|---|------|
| Qy | 61 | GGCTCCATGTTCCGGAGCGAGGAGTGGCCCTGGTCCAGCTCTTTCTGCCCCACAGCGGCT | 120 |
| Db | 61 | | |
| Qy | 121 | GGCTCCATGTTCCGGAGCGAGGAGTGGCCCTGGTCCAGCTCTTTCTGCCCCACAGCGGCT | 120 |
| Db | 121 | | |
| Qy | 121 | GCCTACACCTCGTGAGTTCGGTCGGCCGAGCTGGGCTCGTGAGTTTCAGAGACCTCAAC | 180 |
| Db | 121 | GCCTACACCTCGTGAGTTCGGTCGGCCGAGCTGGGCTCGTGAGTTTCAGAGACCTCAAC | 180 |
| Qy | 181 | GCCTCGGTGAGGCGCTTCAGAGACGCTTTGTGTTGATGTTTGGCGCTGTGAGAGCTG | 240 |
| Db | 181 | GCCTCGGTGAGGCGCTTCAGAGACGCTTTGTGTTGATGTTTGGCGCTGTGAGAGCTG | 240 |
| Qy | 241 | GAGAAAGACTTCACCTTCCTGACGAGAGAGTGGCGGGCTGGGCTGGTCTCCGCCCG | 300 |
| Db | 241 | GAGAAAGACTTCACCTTCCTGACGAGAGAGTGGCGGGCTGGGCTGGTCTCCGCCCG | 300 |
| Qy | 301 | CCAAAGGGGAGGCTGCCGGCACCCACCCCGGGACCTGTGCGCATCCAGAGGAGAGCG | 360 |
| Db | 301 | CCAAAGGGGAGGCTGCCGGCACCCACCCCGGGACCTGTGCGCATCCAGAGGAGAGCG | 360 |
| Qy | 361 | GAGCGCTGGCCACAGGAGCTGGGGATGTGCGGGCAACACAGCAGGCCCTCGCGGGCCAG | 420 |
| Db | 361 | GAGCGCTGGCCACAGGAGCTGGGGATGTGCGGGCAACACAGCAGGCCCTCGCGGGCCAG | 420 |
| Qy | 421 | CTGCACACGCTGCAGCTCCACCGCCGCTGTACGCCAGGSCCATGAACCTCAGCTGGCA | 480 |
| Db | 421 | CTGCACACGCTGCAGCTCCACCGCCGCTGTACGCCAGGSCCATGAACCTCAGCTGGCA | 480 |
| Qy | 481 | GCCGCCACACAGATGGGGCTTCAGAGAGCAGCCCTGTCTCCAGGCCCCCGGGGGCCG | 540 |
| Db | 481 | GCCGCCACACAGATGGGGCTTCAGAGAGCAGCCCTGTCTCCAGGCCCCCGGGGGCCG | 540 |
| Qy | 541 | CACAGGACCTGAGGCTCAACTTTGTGCGAGGTCGTTGAGAGCCCAAGGCCCTTGCC | 600 |
| Db | 541 | CACAGGACCTGAGGCTCAACTTTGTGCGAGGTCGTTGAGAGCCCAAGGCCCTTGCC | 600 |
| Qy | 601 | CTAGAGCCCTGCTCTGAGAGGCTGCGCGGCTTCTCTATTGCCAGTTTCAGGAGCTG | 660 |
| Db | 601 | CTAGAGCCCTGCTCTGAGAGGCTGCGCGGCTTCTCTATTGCCAGTTTCAGGAGCTG | 660 |
| Qy | 661 | GAGCAGCCGCTGGAGCACCCCGTGACGGGCGAGCCAGCCGTGGATGACCTTCTCTCATC | 720 |
| Db | 661 | GAGCAGCCGCTGGAGCACCCCGTGACGGGCGAGCCAGCCGTGGATGACCTTCTCTCATC | 720 |
| Qy | 721 | TCCTACTGGGTGAGCAGATCGGACAGAAATCCGCAAGATCAAGGATGCTTCCATGCG | 780 |
| Db | 721 | TCCTACTGGGTGAGCAGATCGGACAGAAATCCGCAAGATCAAGGATGCTTCCATGCG | 780 |
| Qy | 781 | CACGTCTTCCGTTTCTGACAGGAGGAGCCCGCTCGGGGCCCTGCGAGCAGCTGCAA | 840 |
| Db | 781 | CACGTCTTCCGTTTCTGACAGGAGGAGCCCGCTCGGGGCCCTGCGAGCAGCTGCAA | 840 |
| Qy | 841 | CAGCAGAGCCAGGAGCTGACAGGAGTCTCGGGGAGACAGAGCGGTTCTTGAGCAGGCTG | 900 |
| Db | 841 | CAGCAGAGCCAGGAGCTGACAGGAGTCTCGGGGAGACAGAGCGGTTCTTGAGCAGGCTG | 900 |
| Qy | 901 | CTAGGCGGGTGTGACAGCTGCTGCGCCAGGGCAGGTGCAAGTTCACAGATGAGGCC | 960 |
| Db | 901 | CTAGGCGGGTGTGACAGCTGCTGCGCCAGGGCAGGTGCAAGTTCACAGATGAGGCC | 960 |
| Qy | 961 | GTGTACTGGCCCTGAAACAGTGCAGGTGAGCACACGACACAGTGCCTCATTTGCCGAG | 1020 |
| Db | 961 | GTGTACTGGCCCTGAAACAGTGCAGGTGAGCACACGACACAGTGCCTCATTTGCCGAG | 1020 |
| Qy | 1021 | GCCTGGTGTCTGTGCGAGACCTTCCCGCCCTGTGAGGAGGCCCTTGGCGGACAGCTCGATG | 1080 |
| Db | 1021 | GCCTGGTGTCTGTGCGAGACCTTCCCGCCCTGTGAGGAGGCCCTTGGCGGACAGCTCGATG | 1080 |
| Qy | 1081 | GAGAGGGAGTGTGCTGGCTGACCGCATCCCTTCCCGGACATGCCCCCACTC | 1140 |
| Db | 1081 | GAGAGGGAGTGTGCTGGCTGACCGCATCCCTTCCCGGACATGCCCCCACTC | 1140 |

| | | | | | | |
|----|------|--------------------|-------------------------------|-------------------------------|------------------|------|
| QY | 1141 | ATCGGACCAACCGCTTCA | CGGCCAGCTT | CAGGGCATCGTGGATCGCTACGGCTGGGC | 1200 | |
| DB | 1141 | ATCGGACCAACCGCTTCA | CGGCCAGCTT | CAGGGCATCGTGGATCGCTACGGCTGGGC | 1200 | |
| QY | 1201 | CGCTACGAGAGGTCAAC | CGCGTCCCTAC | CACCATCATCTTCCCTTCTCTGTTTGGCT | 1260 | |
| DB | 1201 | CGCTACGAGAGGTCAAC | CGCGTCCCTAC | CACCATCATCTTCCCTTCTCTGTTTGGCT | 1260 | |
| QY | 1261 | GTGATGTTGGGGATG | GGGCCACGGGCTG | CTCATGTTCTTTCGCTTGGCCATGGTTC | 1320 | |
| DB | 1261 | GTGATGTTGGGGATG | GGGCCACGGGCTG | CTCATGTTCTTTCGCTTGGCCATGGTTC | 1320 | |
| QY | 1321 | CTTTGGGAGAACCGA | CGGCTGTGAACCGCGCAGAACGAGATCTG | GACAGATTTCTTCTTC | 1380 | |
| DB | 1321 | CTTTGGGAGAACCGA | CGGCTGTGAACCGCGCAGAACGAGATCTG | GACAGATTTCTTCTTC | 1380 | |
| QY | 1381 | AGGGGCGGCTACCTG | CTCTGCTTATGGGCTGTTC | CATCTACACCGGCTTCATCTAC | 1440 | |
| DB | 1381 | AGGGGCGGCTACCTG | CTCTGCTTATGGGCTGTTC | CATCTACACCGGCTTCATCTAC | 1440 | |
| QY | 1441 | AACGAGTGCTTCA | GTGCGCCACAGCATCTTCCCTCGGGCTG | GGAGTGTGGCCGCGCATG | 1500 | |
| DB | 1441 | AACGAGTGCTTCA | GTGCGCCACAGCATCTTCCCTCGGGCTG | GGAGTGTGGCCGCGCATG | 1500 | |
| QY | 1501 | GCCAAACGAGTCTG | GGCTGGAGTGATGCA | TTCCTGGCCAGCACACGATGCTTAC | 1560 | |
| DB | 1501 | GCCAAACGAGTCTG | GGAGTGATGCA | TTCCTGGCCAGCACACGATGCTTAC | 1560 | |
| QY | 1561 | CCCAACGTCAAC | CGGCTGTCTTCCCTGGGACCTAC | CCCTTTGGCATCGATCTTATTTGGAGC | 1620 | |
| DB | 1561 | CCCAACGTCAAC | CGGCTGTCTTCCCTGGGACCTAC | CCCTTTGGCATCGATCTTATTTGGAGC | 1620 | |
| QY | 1621 | CTGGCTGCGCAAC | CACTTGAGCTTCTCAA | CTCTTCAAGATGAAGATGTC | CGTCACTCCTG | 1680 |
| DB | 1621 | CTGGCTGCGCAAC | CACTTGAGCTTCTCAA | CTCTTCAAGATGAAGATGTC | CGTCACTCCTG | 1680 |
| QY | 1681 | GGGCTCGTGCA | CATGCGCTTTGGGCTGGTCT | CTGGAGTCTTCAA | ACGAGTGCACTTTGGC | 1740 |
| DB | 1681 | GGGCTCGTGCA | CATGCGCTTTGGGCTGGTCT | CTGGAGTCTTCAA | ACGAGTGCACTTTGGC | 1740 |
| QY | 1741 | CAGAGCACCGGCTG | CTGTGGAGACGCTCGCGGAGCTCA | CTTCCCTGCTGGGACTCTTC | 1800 | |
| DB | 1741 | CAGAGCACCGGCTG | CTGTGGAGACGCTCGCGGAGCTCA | CTTCCCTGCTGGGACTCTTC | 1800 | |
| QY | 1801 | GGTTACTCGTGT | TCTAGTCACTACA | AGTGGCTGTGTCTGGGCTGC | CAGGGCCGCC | 1860 |
| DB | 1801 | GGTTACTCGTGT | TCTAGTCACTACA | AGTGGCTGTGTCTGGGCTGC | CAGGGCCGCC | 1860 |
| QY | 1861 | TGCCCAGCATCTCAT | ATCCATCAACA | CATGTTCTTCTCCACAGCCCGCAGCAAC | 1920 | |
| DB | 1861 | TGCCCAGCATCTCAT | ATCCATCAACA | CATGTTCTTCTCCACAGCCCGCAGCAAC | 1920 | |
| QY | 1921 | AGGCTGCTCTAC | CCCCCGGAGGAGTGGTCCAGGCCAC | CGTGGTGGTCTGGCCTTGGCC | 1980 | |
| DB | 1921 | AGGCTGCTCTAC | CCCCCGGAGGAGTGGTCCAGGCCAC | CGTGGTGGTCTGGCCTTGGCC | 1980 | |
| QY | 1981 | ATGGTGCCCATCT | CTGTGTGGCACAC | CCCTGCACTGTGCA | CCGCGCGC | 2040 |
| DB | 1981 | ATGGTGCCCATCT | CTGTGTGGCACAC | CCCTGCACTGTGCA | CCGCGCGC | 2040 |
| QY | 2041 | CTGCGGAGGAGG | CCGCTGACGACAGGAGGAAAA | CAAGCCGGGTTGCTGGA | CCTGCCT | 2100 |
| DB | 2041 | CTGCGGAGGAGG | CCGCTGACGACAGGAGGAAAA | CAAGCCGGGTTGCTGGA | CCTGCCT | 2100 |
| QY | 2101 | GACGCATCTGTGA | ATGGCTGGAGCTCGATGAGGAA | AAGCAGGGGGCTGGATGAA | 2160 | |
| DB | 2101 | GACGCATCTGTGA | ATGGCTGGAGCTCGATGAGGAA | AAGCAGGGGGCTGGATGAA | 2160 | |
| QY | 2161 | GAGGAGGCGAGCT | CTGTCCTCCGAGGTGCTCAT | GCACAGGCCATCCAC | CACCATCGAG | 2220 |
| DB | 2161 | GAGGAGGCGAGCT | CTGTCCTCCGAGGTGCTCAT | GCACAGGCCATCCAC | CACCATCGAG | 2220 |
| QY | 2221 | TTCTGCTGGGCTG | CGCTCTCCAAACAC | CGCCTCTTACCTGGGCTGTGGGCTT | CGAGCCTG | 2280 |

```
Db 2221 TTCTGCTGGCTGCGTCTCCACACCGCTCTACTCTGCGCTGTGGCCCTGAGCCTG 2280
QY 2281 GCCACGCCAGCTGTCGAGAGTTCTGTGGCCATGATGCGATAGCCCTGGGCCCTG 2340
Db 2281 GCCACGCCAGCTGTCGAGAGTTCTGTGGCCATGATGCGATAGCCCTGGGCCCTG 2340
QY 2341 GGCCTGGAGTGGGCTGGGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 2400
Db 2341 GGCCTGGAGTGGGCTGGGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 2400
QY 2401 ATGACCGTGGCTATCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
Db 2401 ATGACCGTGGCTATCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
QY 2461 CTGCACTGGGTGAATTCAGAACTTCTACTAGGCAAGGCTTCTAGGCAAGGCTT 2520
Db 2461 CTGCACTGGGTGAATTCAGAACTTCTACTAGGCAAGGCTTCTAGGCAAGGCTT 2520
QY 2521 TTCACTTCTGCTGACAGATGACTAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
Db 2521 TTCACTTCTGCTGACAGATGACTAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
QY 2581 GACCTCTGAGGCAAGGAGGATTAAGACGGTCCGCGCTGGCAAAAAAAAAAAAAA 2640
Db 2581 GACCTCTGAGGCAAGGAGGATTAAGACGGTCCGCGCTGGCAAAAAAAAAAAAAA 2640
```

RESULT 4

```
US-09-023-655-916
; Sequence 916, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 916:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2655 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
```

```
; LIBRARY: GENBANK
; CLONE: g1245045
; US-09-023-655-916
```

```
Query Match 100.0%; Score 2640; DB 3; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCTGCGCGAGCGGCGAGCGAGCGGCGGCGAGCACACCCCGGGGACCATG 60
Db 1 CGCGCTGCGCGAGCGGCGAGCGAGCGGCGGCGAGCACACCCCGGGGACCATG 60
QY 61 GGCCTCCTATGTTCCGGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTTGGCCACAGCGGT 120
Db 61 GGCCTCCTATGTTCCGGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTTGGCCACAGCGGT 120
QY 121 GGCCTACACCTGCTGAGTGGGCTGGGCGAGCTGGGCTCGTGGAGTTTACAGACCTCAAC 180
Db 121 GGCCTACACCTGCTGAGTGGGCTGGGCGAGCTGGGCTCGTGGAGTTTACAGACCTCAAC 180
QY 181 GGCCTCGGTGAGCGCTTCCAGAGACGCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Db 181 GGCCTCGGTGAGCGCTTCCAGAGACGCTTTTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
QY 241 GAGAAGACCTTCACTTCCAGAGGAGGTGGCGGCGGCTGGGCTGGTCTGCTGCTGCTGCT 300
Db 241 GAGAAGACCTTCACTTCCAGAGGAGGTGGCGGCGGCTGGGCTGGTCTGCTGCTGCTGCT 300
QY 301 CCAAGGGGAGGCTGCGCGCACCCCGGACCTGCTGCGCATCCAGAGGAGAGCG 360
Db 301 CCAAGGGGAGGCTGCGCGCACCCCGGACCTGCTGCGCATCCAGAGGAGAGCG 360
QY 361 GAGCGCTGCGCCAGAGAGCTGCGGAGATGTGCGGGGAAACAGAGGCGCTTGGCGGCCAG 420
Db 361 GAGCGCTGCGCCAGAGAGCTGCGGAGATGTGCGGGGAAACAGAGGCGCTTGGCGGCCAG 420
QY 421 CTGCAACAGCTGAGCTCCAGCGCGCTGCTACCGCGGCGCATGAACTTCAAGCTGGCA 480
Db 421 CTGCAACAGCTGAGCTCCAGCGCGCTGCTACCGCGGCGCATGAACTTCAAGCTGGCA 480
QY 481 GCGCGCCACACAGATGGGCGCTCAGAGAGGAGCGCCCTGCTCCAGGCGCCCGGGGGCGG 540
Db 481 GCGCGCCACACAGATGGGCGCTCAGAGAGGAGCGCCCTGCTCCAGGCGCCCGGGGGCGG 540
QY 541 CACAGGACCTGAGGCTCACTTTGTGTGAGTGGCGGTGGAGCCCGACAGGCGCCCTGCC 600
Db 541 CACAGGACCTGAGGCTCACTTTGTGTGAGTGGCGGTGGAGCCCGACAGGCGCCCTGCC 600
QY 601 CTAGAGCGCTGCTCTGAGGCGCTTCCCGCGCTTCTCTATTTGCGAGCTTCAAGGAGCTG 660
Db 601 CTAGAGCGCTGCTCTGAGGCGCTTCCCGCGCTTCTCTATTTGCGAGCTTCAAGGAGCTG 660
QY 661 GAGCAGCGCTGAGCAGACCCCGTGTGAGCGGCGAGCGCCAGCGATGACCTTCTCTCATC 720
Db 661 GAGCAGCGCTGAGCAGACCCCGTGTGAGCGGCGAGCGCCAGCGATGACCTTCTCTCATC 720
QY 721 TCCTACTGGGTTGAGCAGATCGGAAGAGATCCGCAAGATCACGAGCTGTCTTCACTGC 780
Db 721 TCCTACTGGGTTGAGCAGATCGGAAGAGATCCGCAAGATCACGAGCTGTCTTCACTGC 780
QY 781 CACGCTTTCCCGTTTCTGAGCAGGAGGCGCCCTCGGGGCGCTGCGAGGCTGCAAA 840
Db 781 CACGCTTTCCCGTTTCTGAGCAGGAGGCGCCCTCGGGGCGCTGCGAGGCTGCAAA 840
QY 841 CAGCAGAGCAGGAGCTGCGAGGAGTCTTCTCGGGGAGACAGAGCGGTTCCTGAGCCAGGTG 900
Db 841 CAGCAGAGCAGGAGCTGCGAGGAGTCTTCTCGGGGAGACAGAGCGGTTCCTGAGCCAGGTG 900
QY 901 CTAGGCGGGTGTGAGGCTGCTGCGCGAGGAGGTGAGGTGCAAGATGAAGGCC 960
Db 901 CTAGGCGGGTGTGAGGCTGCTGCGCGAGGAGGTGAGGTGCAAGATGAAGGCC 960
QY 961 GTGTACCTGGCCCTGAAACAGTGTGAGCGTGGAGCACACAGCAAGTGCCTCATTTGCCGAG 1020
```


| | | | |
|----|------|---|------|
| QY | 61 | GGCTCCATGTTCCGAGCGAGAGGTGGCCCTGGTCCAGCTCTTTCTGCCCCACAGGGCT | 120 |
| Db | 61 | GGCTCCATGTTCCGAGCGAGAGGTGGCCCTGGTCCAGCTCTTTCTGCCCCACAGGGCT | 120 |
| QY | 121 | GCTACACCTGCGTAGTCCGCTGGCGAGCTGGGCTCTGTGGAGTTACAGACCTCAAC | 180 |
| Db | 121 | GCCTACACCTGCGTAGTCCGCTGGCGAGCTGGGCTCTGTGGAGTTACAGACCTCAAC | 180 |
| QY | 181 | GCCTCGGTGAGCGCTTCCAGAGACGCTTTGTGGTTGATGTTTGGCGCTGTGAGGAGCTG | 240 |
| Db | 181 | GCCTCGGTGAGCGCTTCCAGAGACGCTTTGTGGTTGATGTTTGGCGCTGTGAGGAGCTG | 240 |
| QY | 241 | GAGAGACCTTACCTTCTGAGGAGAGGTGGCGGGCTGGGCTGGTCTGCCCGG | 300 |
| Db | 241 | GAGAGACCTTACCTTCTGAGGAGAGGTGGCGGGCTGGGCTGGTCTGCCCGG | 300 |
| QY | 301 | CCAAAGGAGGCTGCGGCAACCCCAACCCCGGACCTTCTGCGCATCCAGGAGGAGCG | 360 |
| Db | 301 | CCAAAGGAGGCTGCGGCAACCCCAACCCCGGACCTTCTGCGCATCCAGGAGGAGCG | 360 |
| QY | 361 | GAGCGCTGGCCAGAGCTGGGGATGTGGGGCAACAGAGGCCCTGGCGGCCAG | 420 |
| Db | 361 | GAGCGCTGGCCAGAGCTGGGGATGTGGGGCAACAGAGGCCCTGGCGGCCAG | 420 |
| QY | 421 | CTGCACAGCTGCAGCTCCAGCGCGCGTGTACGCCAGGGCCATGAACCTCAGCTGGCA | 480 |
| Db | 421 | CTGCACAGCTGCAGCTCCAGCGCGCGTGTACGCCAGGGCCATGAACCTCAGCTGGCA | 480 |
| QY | 481 | GCGGCCACACAGATGGGCTCAGAGAGGACGCGCTCTCCAGGCCCCGGGGGCGG | 540 |
| Db | 481 | GCGGCCACACAGATGGGCTCAGAGAGGACGCGCTCTCCAGGCCCCGGGGGCGG | 540 |
| QY | 541 | CACAGGACCTGAGGCTCAACTTTGTGGAGGTGCGGTGGAGCCCCACAGGCCCCCTGCC | 600 |
| Db | 541 | CACAGGACCTGAGGCTCAACTTTGTGGAGGTGCGGTGGAGCCCCACAGGCCCCCTGCC | 600 |
| QY | 601 | CTAGAGCGCTGCTCTGAGGGCTGCGCGCGCTTCTCATTTGCCAGCTTCAGGAGCTG | 660 |
| Db | 601 | CTAGAGCGCTGCTCTGAGGGCTGCGCGCGCTTCTCATTTGCCAGCTTCAGGAGCTG | 660 |
| QY | 661 | GAGCAGCGCTGAGACACCCCGTGACGGCGAGCCAGCCAGCTGGATGACCTTCTCATC | 720 |
| Db | 661 | GAGCAGCGCTGAGACACCCCGTGACGGCGAGCCAGCCAGCTGGATGACCTTCTCATC | 720 |
| QY | 721 | TCTTACTGGGTGAGCAGATCCGACAGAGATCCGCAAGATCACCGACTGCTTCACTGC | 780 |
| Db | 721 | TCTTACTGGGTGAGCAGATCCGACAGAGATCCGCAAGATCACCGACTGCTTCACTGC | 780 |
| QY | 781 | CACGTCTTCCCGTTCTGACAGAGAGGAGGCGCGCTCGGGCCCTGACGAGCTGCAAA | 840 |
| Db | 781 | CACGTCTTCCCGTTCTGACAGAGAGGAGGCGCGCTCGGGCCCTGACGAGCTGCAAA | 840 |
| QY | 841 | CAGCAGACCGAGAGCTGAGAGGTCTTCCGGGAGACAGCGGTTCTTGAGCCAGGTG | 900 |
| Db | 841 | CAGCAGACCGAGAGCTGAGAGGTCTTCCGGGAGACAGCGGTTCTTGAGCCAGGTG | 900 |
| QY | 901 | CTAGGCGGGTGTGAGCTGCTGCCCGCAGGGCAGGTCCAGAGATGAAGGCC | 960 |
| Db | 901 | CTAGGCGGGTGTGAGCTGCTGCCCGCAGGGCAGGTCCAGAGATGAAGGCC | 960 |
| QY | 961 | GTGTACTTCCCGCTGAAACAGTGCAGCTGAGCACCAGCACAAGTGCTCATTTGCCGAG | 1020 |
| Db | 961 | GTGTACTTCCCGCTGAAACAGTGCAGCTGAGCACCAGCACAAGTGCTCATTTGCCGAG | 1020 |
| QY | 1021 | GCTGTGTCTGTGCGAGCTTGGCCCGCTTGCAGGAGGCCCTTGGGAGCAGCTCGATG | 1080 |
| Db | 1021 | GCTGTGTCTGTGCGAGCTTGGCCCGCTTGCAGGAGGCCCTTGGGAGCAGCTCGATG | 1080 |
| QY | 1081 | GAGGAGGAGTGAGTCCCTGGCTCACCGCATCCCTTGGCGGGACATGCCCCCCACATC | 1140 |
| Db | 1081 | GAGGAGGAGTGAGTCCCTGGCTCACCGCATCCCTTGGCGGGACATGCCCCCCACATC | 1140 |

| | | | |
|----|------|--|------|
| QY | 1141 | ATCCGACAAACCGCTTCAAGCGCCAGCTTTCAGGGCATCTGTGGATCGCTACGGCGTGGC | 1200 |
| Db | 1141 | ATCCGACAAACCGCTTCAAGCGCCAGCTTTCAGGGCATCTGTGGATCGCTACGGCGTGGC | 1200 |
| QY | 1201 | CGCTACACGAGGTCAACCCCGCTCTTACACCATCATCACCTTCCCTTCTGTTGCT | 1260 |
| Db | 1201 | CGCTACACGAGGTCAACCCCGCTCTTACACCATCATCACCTTCCCTTCTGTTGCT | 1260 |
| QY | 1261 | GTGATGTTGGGGATGTGGGCAACCGGCTGTCTTCTTCTTCCCTTGGCCATGCTC | 1320 |
| Db | 1261 | GTGATGTTGGGGATGTGGGCAACCGGCTGTCTTCTTCTTCCCTTGGCCATGCTC | 1320 |
| QY | 1321 | CTTTGGGAGAACCGACCGGCTGTGAAAGCGCGAGAACGAGATCTGCGAGCTTCTTC | 1380 |
| Db | 1321 | CTTTGGGAGAACCGACCGGCTGTGAAAGCGCGAGAACGAGATCTGCGAGCTTCTTC | 1380 |
| QY | 1381 | AGGGCCGCTACCTGCTCTCTTATAGGGCTGTCTTCTCCATCTACACGGCTTCATCTAC | 1440 |
| Db | 1381 | AGGGCCGCTACCTGCTCTCTTATAGGGCTGTCTTCTCCATCTACACGGCTTCATCTAC | 1440 |
| QY | 1441 | AACGAGTGTTCAGTCCGCCACACAGCATCTTCCCTCGGGCTGAGTGTGGCCCGCATG | 1500 |
| Db | 1441 | AACGAGTGTTCAGTCCGCCACACAGCATCTTCCCTCGGGCTGAGTGTGGCCCGCATG | 1500 |
| QY | 1501 | GCCAAACAGTCTGGCTGGAGTGATGCAATCTTGGGCCACAGCACGATGCTTACCTGGAT | 1560 |
| Db | 1501 | GCCAAACAGTCTGGCTGGAGTGATGCAATCTTGGGCCACAGCACGATGCTTACCTGGAT | 1560 |
| QY | 1561 | CCCAACAGTCTGGCTGGAGTGTCTTCTGGGACCTTACCCCTTGGCATCGATCTTGGAGC | 1620 |
| Db | 1561 | CCCAACAGTCTGGCTGGAGTGTCTTCTGGGACCTTACCCCTTGGCATCGATCTTGGAGC | 1620 |
| QY | 1621 | CTGGCTGCGCAACCACTTGAAGTCTTCACTTCAAGTCAAGATGAGATGTCGTCACTG | 1680 |
| Db | 1621 | CTGGCTGCGCAACCACTTGAAGTCTTCACTTCAAGTCAAGATGAGATGTCGTCACTG | 1680 |
| QY | 1681 | GCGCTGCTGCAATGGGCTTGGGCTGCTTGGAGTCTTCAACCAAGTGCATTTGGC | 1740 |
| Db | 1681 | GCGCTGCTGCAATGGGCTTGGGCTGCTTGGAGTCTTCAACCAAGTGCATTTGGC | 1740 |
| QY | 1741 | CAGAGCACCGGCTGCTGAGAGCGCTGCGGAGCTCACCTTCTTCTGCGGACTCTTC | 1800 |
| Db | 1741 | CAGAGCACCGGCTGCTGAGAGCGCTGCGGAGCTCACCTTCTTCTGCGGACTCTTC | 1800 |
| QY | 1801 | GCTTACTCTGTTCTTAGTCACTACAGTGGCTGTGTGTGGCTGCGGCGCGCC | 1860 |
| Db | 1801 | GCTTACTCTGTTCTTAGTCACTACAGTGGCTGTGTGTGGCTGCGGCGCGCC | 1860 |
| QY | 1861 | TCGCCAGCATCTCATCACTTCAATCAATGTTCTTCTTCTCCACAGCCCGCAGCAAC | 1920 |
| Db | 1861 | TCGCCAGCATCTCATCACTTCAATCAATGTTCTTCTTCTCCACAGCCCGCAGCAAC | 1920 |
| QY | 1921 | AGGCTGCTTACCCCGCAGAGGTGTTCAGGCCACGCTGGTGGTCTTGGCTTGGCC | 1980 |
| Db | 1921 | AGGCTGCTTACCCCGCAGAGGTGTTCAGGCCACGCTGGTGGTCTTGGCTTGGCC | 1980 |
| QY | 1981 | ATGGTCCCATCTCTGCTTGGCACACCTTGCACCTGCTGTCACCGCCACCGCCCGC | 2040 |
| Db | 1981 | ATGGTCCCATCTCTGCTTGGCACACCTTGCACCTGCTGTCACCGCCACCGCCCGC | 2040 |
| QY | 2041 | CTGCGGAGGAGCGCGCTGACCGACAGGAGGAAACAGGGCGGGTGTGAGACCTGCT | 2100 |
| Db | 2041 | CTGCGGAGGAGCGCGCTGACCGACAGGAGGAAACAGGGCGGGTGTGAGACCTGCT | 2100 |
| QY | 2101 | GACGATCTGTGAATGCTGAGTCTCCATGAGGAAAGGAGGCGGGCTTGAATGAA | 2160 |
| Db | 2101 | GACGATCTGTGAATGCTGAGTCTCCATGAGGAAAGGAGGCGGGCTTGAATGAA | 2160 |
| QY | 2161 | GAGGAGGAGCTGCTCCCTCGAGGTGTCTATGCAACCGGCTTCAACACCATCGAG | 2220 |
| Db | 2161 | GAGGAGGAGCTGCTCCCTCGAGGTGTCTATGCAACCGGCTTCAACACCATCGAG | 2220 |
| QY | 2221 | TTCCTGCTGGGCTGCTCTCCAAACCGCCCTCTACCTGCGCTCTGTGGGCTTGGGCTG | 2280 |

| | | | |
|----------------|------|---|------|
| D _b | 2221 | TTCGCTCGGTGGTGGTCTCCAAACCGCCTCTACTCGCGCTGTGGGCCCTTAGCCTG | 2280 |
| Q _y | 2281 | GCCACGCCAGCTGTCGAGGTTCTGTGGGCCATGGTGATGGCATAGGCCTTGGGCCCTG | 2340 |
| D _b | 2291 | GCCACGCCAGCTGTCGAGGTTCTGTGGGCCATGGTGATGGCATAGGCCTTGGGCCCTG | 2340 |
| Q _y | 2341 | GGCCGGAGGTGGGCGTGGCGGCTGTGGTGCTGGTCCCCATCTTTGGCGCCTTTGGCCGTG | 2400 |
| D _b | 2341 | GGCCGGAGGTGGGCGTGGCGGCTGTGGTGCTGGTCCCCATCTTTGGCGCCTTTGGCCGTG | 2400 |
| Q _y | 2401 | ATGACCGTGGCTATCCTGCTGGTGATGGAGGGACTCTCAGCCCTTCCTGCACGCGCTTCGCG | 2460 |
| D _b | 2401 | ATGACCGTGGCTATCCTGCTGGTGATGGAGGGACTCTCAGCCCTTCCTGCACGCGCTTCGCG | 2460 |
| Q _y | 2461 | CTGCACCTGGGTGGAAATTCAGAAACAAGTTCTACTCAGGCA CGGGCTACAAGCTGAGTCCC | 2520 |
| D _b | 2461 | CTGCACCTGGGTGGAAATTCAGAAACAAGTTCTACTCAGGCA CGGGCTACAAGCTGAGTCCC | 2520 |
| Q _y | 2521 | TTACCTTCGTGCGCACAGATGACTTAGGGCCCACTGCAGGTCTGCGCAGACCTCTTCCT | 2580 |
| D _b | 2521 | TTACCTTCGTGCGCACAGATGACTTAGGGCCCACTGCAGGTCTGCGCAGACCTCTTCCT | 2580 |
| Q _y | 2581 | GACCTCTCAGGCAGAGAGGAATAAGACGCTCCGCCCTGGCAAAAAAAAAAAAAA | 2640 |
| D _b | 2581 | GACCTCTCAGGCAGAGAGGAATAAGACGCTCCGCCCTGGCAAAAAAAAAAAAAA | 2640 |

RESULT 6
US-09-949-016-2404
; Sequence 2404, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2404
; LENGTH: 2654
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2404

| | Query Match | 81.1% | Score 2142; | DB 3; | Length 2654; |
|----|-----------------------|--|---------------|-----------|--------------|
| | Best Local Similarity | 99.7%; | Pred. No. 0; | | |
| | Matches 2582; | Conservative 0; | Mismatches 4; | Indels 3; | Gaps 1; |
| Qy | 38 | GCGCAGACACACCCGGGACCATGGCTCCATGTTCCGAGCGAGAGCTGGCCCTGCTGC | 97 | | |
| Db | 66 | GGCGAGACACCCGGGAGCCATGGCTCCATGTTCCGAGCGAGAGCTGGCCCTGCTGC | 125 | | |
| Qy | 98 | AGCTCTTTTGTGCCACACGCGCTGCCCTACACCTCGGTGAGTTCGGCTCGGCGAGCTGGGCC | 157 | | |
| Db | 126 | AGCTCTTTTGTGCCACACGCGCTGCCCTACACCTCGGTGAGTTCGGCTCGGCGAGCTGGGCC | 185 | | |
| Qy | 158 | TCGTGGAGTTTCAGAGACTCAACGCCCTCGGTGAGCGCCTTCCAGAGACGCTTTGTGCTGTG | 217 | | |
| Db | 186 | TCGTGGAGTTTCAGAGACTCAACGCCCTCGGTGAGCGCCTTCCAGAGACGCTTTGTGCTGTG | 245 | | |
| Qy | 218 | ATGTTTGGCGCTGTGAGAGCTGGAGAAGACCTTTCACCTTCCTGCAGAGAGAGTTCGGC | 277 | | |
| Db | 246 | ATGTTTCGGCGCTGTGAGAGCTGGAGAAGACCTTTCACCTTCCTGCAGAGAGAGTTCGGC | 305 | | |

| | | | |
|----|------|--|------|
| QY | 278 | GGGCTGGGCTGGTCTCTGCCCCCGCCAAAGGGGAGGCTGCGGCACCCCCACCCCGGGACC | 337 |
| Db | 306 | GGGCTGGGCTGGTCTCTGCCCCCGCCAAAGGGGAGGCTGCGGCACCCCCACCCCGGGACC | 365 |
| QY | 338 | TGCTGGCANTCCAGGAGGAGACGGAGCGCTTGCCCCCAGGAGCTGCGGGATGTGCGGGGCA | 397 |
| Db | 366 | TGCTGGCANTCCAGGAGGAGACGGAGCGCTTGCCCCCAGGAGCTGCGGGATGTGCGGGGCA | 425 |
| QY | 398 | ACCAGACGGCCCTTCGGGCCCCAGCTGCACACAGCTGCAGCTCCACCGCCCGTGTACGCC | 457 |
| Db | 426 | ACCAGACGGCCCTTCGGGCCCCAGCTGCACACAGCTGCAGCTCCACCGCCCGTGTACGCC | 485 |
| QY | 458 | AGGGCCATGAACCTCAGCTGGCAGCCGCCACACAGATGGGGCTCAGAGAGACGCCCC | 517 |
| Db | 486 | AGGGCCATGAACCTCAGCTGGCAGCCGCCACACAGATGGGGCTCAGAGAGACGCCCC | 545 |
| QY | 518 | TGCTCCAGGGCCCCCGGGGGCCGACACAGAGACCTGAGGGTCAACTTTGTGGCAGGTGCCG | 577 |
| Db | 546 | TGCTCCAGGGCCCCCGGGGGCCGACACAGAGACCTGAGGGTCAACTTTGTGGCAGGTGCCG | 605 |
| QY | 578 | TGGAGCCCCCAAGCGCCCTGCCCTAGAGCGCTGTCTCTGGAGGGCTGCGCGGGTTCC | 637 |
| Db | 606 | TGGAGCCCCCAAGCGCCCTGCCCTAGAGCGCTGTCTCTGGAGGGCTGCGCGGGTTCC | 665 |
| QY | 638 | TCATTGCCAGTTTCAGGGAGCTGAGAGCAGCGCTGAGACACCCCGTGA CGGGCGAGCCAG | 697 |
| Db | 666 | TCATTGCCAGTTTCAGGGAGCTGAGAGCAGCGCTGAGACACCCCGTGA CGGGCGAGCCAG | 725 |
| QY | 698 | CCAGTGGATGACCTTCCTCATCTCTACTTGGGGTGAGCAGATCGGACAGAAATCCGCA | 757 |
| Db | 726 | CCAGTGGATGACCTTCCTCATCTCTACTTGGGGTGAGCAGATCGGACAGAAATCCGCA | 785 |
| QY | 758 | AGATTCAAGGACTGTCTTCACTTGCCACAGCTTCCCGTTTCTGCAGCAGGAGGAGCGCCGCC | 817 |
| Db | 786 | AGATTCAAGGACTGTCTTCACTTGCCACAGCTTCCCGTTTCTGCAGCAGGAGGAGCGCCGCC | 845 |
| QY | 818 | TCGGGGCCCTGCAGCAGCTGCAACAGCAGAGCCAGGAGTGCAGGAGGTCTCTGGGGAGA | 877 |
| Db | 846 | TCGGGGCCCTGCAGCAGCTGCAACAGCAGAGCCAGGAGTGCAGGAGGTCTCTGGGGAGA | 905 |
| QY | 878 | CAGAGCGGTTCTTGAGCCAGGTCTAGGCGGGTGCTGCAGCTGCTGCCGCCAGGGCAGG | 937 |
| Db | 906 | CAGAGCGGTTCTTGAGCCAGGTCTAGGCGGGTGCTGCAGCTGCTGCCGCCAGGGCAGG | 965 |
| QY | 938 | TGCAGGTCCACAAGATGAAGCCGTGTACTTGGCCCTTGAAACAGTGCAGCGTGAGCACCA | 997 |
| Db | 966 | TGCAGGTCCACAAGATGAAGCCGTGTACTTGGCCCTTGAAACAGTGCAGCGTGAGCACCA | 1025 |
| QY | 998 | CGCAAGATGCTCATTTGCCAGGCTTGTGCTCTGTGCGAGACCTGCCCGCCCTGCAGG | 1057 |
| Db | 1026 | CGCAAGATGCTCATTTGCCAGGCTTGTGCTCTGTGCGAGACCTGCCCGCCCTGCAGG | 1085 |
| QY | 1058 | AGGCCCTGCGGGACAGCTCGATGGAGGAGGAGTGAGTGCCGTGGCTCACCGCATCCCT | 1117 |
| Db | 1086 | AGGCCCTGCGGGACAGCTCGATGGAGGAGGAGTGAGTGCCGTGGCTCACCGCATCCCT | 1145 |
| QY | 1118 | GCCGGGACATGCCCCCACTCATTCGCAACCAACCGCTTCACGGCCAGCTTCCAGGGCA | 1177 |
| Db | 1146 | GCCGGGACATGCCCCCACTCATTCGCAACCAACCGCTTCACGGCCAGCTTCCAGGGCA | 1205 |
| QY | 1178 | TCGTGGATTCGCTACGGCGTGGGCGCTTACAGAGAGTCAACCCCGCTGCCCTACCATCA | 1237 |
| Db | 1206 | TCGTGGATTCGCTACGGCGTGGGCGCTTACAGAGAGTCAACCCCGCTGCCCTACCATCA | 1265 |
| QY | 1238 | TCACCTTCCCTTCCCTGTTGCTGTGATGTTTCGGGATGTGGGCGCAGGGCTGCTCATGT | 1297 |
| Db | 1266 | TCACCTTCCCTTCCCTGTTGCTGTGATGTTTCGGGATGTGGGCGCAGGGCTGCTCATGT | 1325 |
| QY | 1298 | TCCTTCTTCCCTTCCCTGCTTCTTCGCGAGAACCGACCGGCTGTGAAAGCCGCGCAGA | 1357 |
| Db | 1326 | TCCTTCTTCCCTTCCCTGCTTCTTCGCGAGAACCGACCGGCTGTGAAAGCCGCGCAGA | 1385 |
| QY | 1358 | ACGAGATCTGCGACACTTCTTCAGGGGCGCGCTACCTGCTCTCTGCTTTATGGGCTGTCT | 1417 |

1386 ACAGATCTGGCAGACTTTCTTTCAGGGCGCTACCTGCTCTGCTTATGGGCTGTCT 1445
1418 CCATCTACACCGGCTTCTATCAACAGAGTGCTTCAAGTCGGCCACACAGATCTTCCCT 1477
1446 CCATCTACACCGGCTTCTATCAACAGAGTGCTTCAAGTCGGCCACACAGATCTTCCCT 1505
1478 CGGGCTGGAGTGTGGCGGCATGGCCAAACAGTCTGGCTGGAGTGAATGATCTTCTGGCCC 1537
1506 CGGGCTGGAGTGTGGCGGCATGGCCAAACAGTCTGGCTGGAGTGAATGATCTTCTGGCCC 1565
1538 AGCACACGATGCTTACCTTGGATCCCAAGGTCAACCGGTGTCTTCTGGGACCTTACCTCT 1597
1566 AGCACACGATGCTTACCTTGGATCCCAAGGTCAACCGGTGTCTTCTGGGACCTTACCTCT 1625
1598 TTGGCATCATCTTATTTGGAGGCTGGCTGGCCAAACACATGAGCTTCTCAATCTCTTCA 1657
1626 TTGGCATCATCTTATTTGGAGGCTGGCTGGCCAAACACATGAGCTTCTCAATCTCTTCA 1685
1658 AGATGAAGATGTCGGTCACTCTGGGCGTGTGCACATGGCTTTGGGGTGGTCTCGGAG 1717
1686 AGATGAAGATGTCGGTCACTCTGGGCGTGTGCACATGGCTTTGGGGTGGTCTCGGAG 1745
1718 TCTTCAACACGCTGCACTTTGGCCAGAGGACCGGCTGCTGTGGAGACGCTGCCGGAGC 1777
1746 TCTTCAACACGCTGCACTTTGGCCAGAGGACCGGCTGCTGTGGAGACGCTGCCGGAGC 1805
1778 TCACCTTCTCTGGGACCTTTCGGTTACTGCTGTGTTCTAGTCACTCAAGTGGCTGT 1837
1806 TCACCTTCTCTGGGACCTTTCGGTTACTGCTGTGTTCTAGTCACTCAAGTGGCTGT 1865
1838 GTGTCTGGGCTGCAGGGCGGCTCG- --CCAGCATCTCTCACTTCACTCAACATGT 1894
1866 GTGTCTGGGCTGCAGGGCGGCTTCGGCCCGCAGCATCTCTCACTTCACTCAACATGT 1925
1895 TCTCTTCTCCACAGCCCCAGCAACAGGCTGCTCTTACCCCCGGCAGGAGGTGGTCCAGG 1954
1926 TCTCTTCTCCACAGCCCCAGCAACAGGCTGCTCTTACCCCCGGCAGGAGGTGGTCCAGG 1985
1955 CCAGCTGTGTGCTTGGGCTTGGCCATGGTGCCCATCTGTCTGTGGCACACCCCTGC 2014
1986 CCAGCTGTGTGCTTGGGCTTGGCCATGGTGCCCATCTGTCTGTGGCACACCCCTGC 2045
2015 ACTGTGTGACCCGACCGCGGCTGCGAGAGGCGGCTGACCCGACAGGAGGAAA 2074
2046 ACTGTGTGACCGCCACCGCGGCTGCGAGAGGCGGCTGACCCGACAGGAGGAAA 2105
2075 ACAAGCCGGGTTGTGGACCTTGCCTGACGCACTCTGTGAATGGCTGGAGCTCCGATGAGG 2134
2106 ACAAGCCGGGTTGTGGACCTTGCCTGACGCACTCTGTGAATGGCTGGAGCTCCGATGAGG 2165
2135 AAAAGCCAGGGGCTTGGATGATGAAGAGAGGCGGAGCTGCTCCCTCGAGGTGCTCA 2194
2166 AAAAGCCAGGGGCTTGGATGATGAAGAGAGGCGGAGCTGCTCCCTCGAGGTGCTCA 2225
2195 TGACACAGGCGCATCCACACATCTGAGTCTGTGGGCTGCGTCTCAACACCGCTCTCT 2254
2226 TGACACAGGCGCATCCACACATCTGAGTCTGTGGGCTGCGTCTCAACACCGCTCTCT 2285
2255 ACTGCGCTGTGGGCGCTTGAAGCTTGGCCCAAGCCAGCTGTCCGAGGTCTGTGGGCA 2314
2286 ACTGCGCTGTGGGCGCTTGAAGCTTGGCCCAAGCCAGCTGTCCGAGGTCTGTGGGCA 2345
2315 TGGTGAATGGCGATAGGCTTGGGCTTGGGCGGAGGTGGGCTGGGCTGTGGTGTCTGG 2374
2346 TGGTGAATGGCGATAGGCTTGGGCTTGGGCGGAGGTGGGCTGGGCTGTGGTGTCTGG 2405
2375 TCCCATCTTTTGGCGCTTGGCGTGAATGACCGTGGCTATCTGTGCTGGTGGAGGGAC 2434
2406 TCCCATCTTTTGGCGCTTGGCGTGAATGACCGTGGCTATCTGTGCTGGTGGAGGGAC 2465
2435 TCTAGCCTTCTTGACGCGCTTGGGCTGCACTGGGTTGGAATTCAGAGAACTTCTACT 2494

Db 2466 TCTCAGCCTTCTCTGCAACCGCTGCGGCTGCACTGGGTGGAATTCAGAACTTCTACT 2525
QY 2495 CAGGCAACGGGCTAAGCTGAGTCCCTTCACTTGGCTGCCACAGATGATGAGGCCCCAC 2554
Db 2526 CAGGCAACGGGCTAAGCTGAGTCCCTTCACTTGGCTGCCACAGATGATGAGGCCCCAC 2585
QY 2555 TGCAGGCTCTGCGACAGCTCTTCTGACCTCTCTGAGGAGGAGGAGAAATAAGACGGTCC 2614
Db 2586 TGCAGGCTCTGCGACAGCTCTTCTGACCTCTCTGAGGAGGAGGAGAAATAAGACGGTCC 2645
QY 2615 GCCCTGGCA 2623
Db 2646 GCCCTGGCA 2654

RESULT 7

US-09-949-016-2405
; Sequence 2405, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2405
; LENGTH: 2654
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2405

Query Match 81.1%; Score 2142; DB 3; Length 2654;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2582; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 38 GCGAGCACACCCGGGACCATGGGCTCCATGTTCCGAGCGAGAGGTGGCCCTGGTCC 97
Db 66 GCGAGCACACCCGGGACCATGGGCTCCATGTTCCGAGCGAGAGGTGGCCCTGGTCC 125
QY 98 AGCTCTTTCTGCCACACGCGCTGCTACCTGCTGAGTCGGCTGGCGAGAGTGGGCG 157
Db 126 AGCTCTTTCTGCCACACGCGCTGCTACCTGCTGAGTCGGCTGGCGAGAGTGGGCG 185
QY 158 TCGTGGAGTTCAGAGACCTCAACGCTCGGTGAGCGCTTCAGAGACGCTTTGTGTTG 217
Db 186 TCGTGGAGTTCAGAGACCTCAACGCTCGGTGAGCGCTTCAGAGACGCTTTGTGTTG 245
QY 218 ATGTTTGGGCTGTGAGAGGCTGGAGAGACCTTCACTTCTCTGAGGAGAGTGGCG 277
Db 246 ATGTTTGGGCTGTGAGAGGCTGGAGAGACCTTCACTTCTCTGAGGAGAGTGGCG 305
QY 278 GGGCTGGGCTGTCTTCTGCCCCCGCCAAAGGGGAGGCTGCGCGCACCCACCCCGGAC 337
Db 306 GGGCTGGGCTGTCTTCTGCCCCCGCCAAAGGGGAGGCTGCGCGCACCCACCCCGGAC 365
QY 338 TGCTCGCATTCAGAGAGAGACGAGCGCTTGCGCCACAGAGCTGCGGAGATGTCGGGCA 397
Db 366 TGCTCGCATTCAGAGAGAGACGAGCGCTTGCGCCACAGAGCTGCGGAGATGTCGGGCA 425
QY 398 ACCACAGGCGCTTGGGCGCCAGCTGACAGCTCAGCTCCAGCCCGCGTGTACGCC 457
Db 426 ACCACAGGCGCTTGGGCGCCAGCTGACAGCTCAGCTCCAGCCCGCGTGTACGCC 485
QY 458 AGGCCATGAACCTCAGCTGGCAGCGCCGCCACACAGATGGGGCTTCAGAGAGACGCC 517

RESULT 8

US-949-016-2975
Sequence 2975, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2975
LENGTH: 2706
TYPE: DNA
ORGANISM: Human
US-949-016-2975

Query Match 80.6%; Score 2127; DB 3; Length 2706;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2567; Conservative 0; Mismatches 4; Indels 3; Gaps 1;
QY 53 GGACCATGGGCTCCATGTTCCGAGCGAGGAGTGGCCCTGGTCAGAGCTTTTCGGCCA 112
DB 133 GGACCATGGGCTCCATGTTCCGAGCGAGGAGTGGCCCTGGTCAGAGCTTTTCGGCCA 192
QY 113 CAGGGCTGCTTACACCTCGTGAGTCGGCTGGGAGAGCTGGGCTCGTGGAGTTCAAG 172
DB 193 CAGGGCTGCTTACACCTCGTGAGTCGGCTGGGAGAGCTGGGCTCGTGGAGTTCAAG 252
QY 173 ACCTCAACGCTCGTGAGCGCTTCCAGAGAGCTTTGTGTTGATGTTTGGCGCTGTG 232
DB 253 ACCTCAACGCTCGTGAGCGCTTCCAGAGAGCTTTGTGTTGATGTTTGGCGCTGTG 312
QY 233 AGGAGCTGAGAGAGCTTCACTTCTCGAGGAGAGTGGCGGGCTGGGCTGGTCTC 292
DB 313 AGGAGCTGAGAGAGCTTCACTTCTCGAGGAGAGTGGCGGGCTGGGCTGGTCTC 372
QY 293 TGCCCCCGCCAAAGGGAGGCTGCGGGCAACCCCAACCCCGGACCTGTGCGATCCAGG 352
DB 373 TGCCCCCGCCAAAGGGAGGCTGCGGGCAACCCCAACCCCGGACCTGTGCGATCCAGG 432
QY 353 AGGAGAGGAGCGCTGGCCCAAGGAGCTGCGGGATGTGGGGCAACCAAGAGGCCCTGC 412
DB 433 AGGAGAGGAGCGCTGGCCCAAGGAGCTGCGGGATGTGGGGCAACCAAGAGGCCCTGC 492
QY 413 GGGCCAGCTGCACCAAGCTGAGCTCCAGCGCCCGTGTCTAGCGAGGAGGATGAACCTC 472
DB 493 GGGCCAGCTGCACCAAGCTGAGCTCCAGCGCCCGTGTCTAGCGAGGAGGATGAACCTC 552
QY 473 AGCTGGAGCGCCCAACAGATGGGGCTCAGAGAGAGCGCCCTGTCTCAGAGGCCCGG 532
DB 553 AGCTGGAGCGCCCAACAGATGGGGCTCAGAGAGAGCGCCCTGTCTCAGAGGCCCGG 612
QY 533 GGGGGCCGACCAAGAGCTGAGGCTCACTTTGTGGAGTGGCGGTCAGAGGCCCAAGG 592
DB 613 GGGGGCCGACCAAGAGCTGAGGCTCACTTTGTGGAGTGGCGGTCAGAGGCCCAAGG 672
QY 593 CCCTGCTCCTAGAGGCTGTCTGAGGGGCTTCCAGAGGCTTCTCATTGCGAGCTTCA 652
DB 673 CCCTGCTCCTAGAGGCTGTCTGAGGGGCTTCCAGAGGCTTCTCATTGCGAGCTTCA 732
QY 653 GGGAGCTGAGAGCGCTGAGAGCAACCCGCTGAGCGGGAGCGCAGCGAGTGGATGACTT 712
DB 732 GGGAGCTGAGAGCGCTGAGAGCAACCCGCTGAGCGGGAGCGCAGCGAGTGGATGACTT 792

DB 733 GGGAGCTGAGAGCGCTGAGAGCAACCCGCTGAGCGGGAGCGCAGCGTGGATGACCT 792
QY 713 TCCTCATCTCTACTTGGGGTGAAGATCGGACAGAGATCGGAGATCGGAGTGTCT 772
DB 793 TCCTCATCTCTACTTGGGGTGAAGATCGGACAGAGATCGGAGATCGGAGTGTCT 852
QY 773 TCCTCATCTCTACTTGGGGTGAAGATCGGACAGAGATCGGAGATCGGAGTGTCT 832
DB 853 TCCTCATCTCTACTTGGGGTGAAGATCGGACAGAGATCGGAGATCGGAGTGTCT 912
QY 833 AGCTCAACAGAGAGCGAGGAGTCTCGGGGAGAGAGAGTCTCGGGGAGAGAGTCTCG 892
DB 913 AGCTCAACAGAGAGCGAGGAGTCTCGGGGAGAGAGTCTCGGGGAGAGAGTCTCG 972
QY 893 GCCAGTGTAGGCGGGGCTGTCAGCTGCTGCGCCAGAGGAGTGTGAGGTCACCAAG 952
DB 973 GCCAGTGTAGGCGGGGCTGTCAGCTGCTGCGCCAGAGGAGTGTGAGGTCACCAAG 1032
QY 953 TGAAGGCGGTGTACCTGGCCCTGAAACAGTGTGAGCTGAGCAGCAGCAGCAAGTGTCTCA 1012
DB 1033 TGAAGGCGGTGTACCTGGCCCTGAAACAGTGTGAGCAGCAGCAGCAGCAAGTGTCTCA 1092
QY 1013 TTGCGAGGCGGTGTGCTGTCGAGACCTGCGCCGCTGTCAGAGAGCGCTTCGCGGACA 1072
DB 1093 TTGCGAGGCGGTGTGCTGTCGAGACCTGCGCCGCTGTCAGAGAGCGCTTCGCGGACA 1152
QY 1073 GCTCATGAGAGGAGTGTGCTGCGCTGTCAGCGCATCCCTGCGGGAGATGCCCT 1132
DB 1153 GCTCATGAGAGGAGTGTGCTGCGCTGTCAGCGCATCCCTGCGGGAGATGCCCT 1212
QY 1133 CCACACTCATCGCACCAACCGCTTCAAGGCGAGCTTCAAGGCGATCGTGGATCGTAGC 1192
DB 1213 CCACACTCATCGCACCAACCGCTTCAAGGCGAGCTTCAAGGCGATCGTGGATCGTAGC 1272
QY 1193 GGTGGGCGGTGTACAGAGAGTCAACCGCGCTCCCTACACCATCATCACTTCCCTTCC 1252
DB 1273 GGTGGGCGGTGTACAGAGAGTCAACCGCGCTCCCTACACCATCATCACTTCCCTTCC 1332
QY 1253 TGTTGTGCTGATGTTTGGGGATGTGGGCGATGTGGGCGATGTGGGCTGTCTTCTTGGCGCTGG 1312
DB 1333 TGTTGTGCTGATGTTTGGGGATGTGGGCGATGTGGGCGATGTGGGCTGTCTTCTTGGCGCTGG 1392
QY 1313 CCATGCTCTTTCGGAGAAACCGACCGGCTGTGAAGAGCGCGAGAAACAGAGATCTGGCAGA 1372
DB 1393 CCATGCTCTTTCGGAGAAACCGACCGGCTGTGAAGAGCGCGAGAAACAGAGATCTGGCAGA 1452
QY 1373 CTCTTCTTCAAGGGCGGTGTACCTGTCTCTGCTTATGCGGCTGTGTCTCACTACCGGCT 1432
DB 1453 CTCTTCTTCAAGGGCGGTGTACCTGTCTCTGCTTATGCGGCTGTGTCTCACTACCGGCT 1512
QY 1433 TCATCTCAACAGAGTGTTCAGTGTGGCGCACCGAGCATCTTCCCTGCGGCTGGAGTGTGG 1492
DB 1513 TCATCTCAACAGAGTGTTCAGTGTGGCGCACCGAGCATCTTCCCTGCGGCTGGAGTGTGG 1572
QY 1493 CCGCATGCGCCAAACAGTCTGGCTGGAGTGTGATTCCTGCGCCAGCAGACAGATGCTTCA 1552
DB 1573 CCGCATGCGCCAAACAGTCTGGCTGGAGTGTGATTCCTGCGCCAGCAGACAGATGCTTCA 1632
QY 1553 CCCTGGATCCCAACGTCAACCGGTGTCTTCTGGGAGCCCTACCCCTTTGGGATCGATCTCA 1612
DB 1633 CCCTGGATCCCAACGTCAACCGGTGTCTTCTGGGAGCCCTACCCCTTTGGGATCGATCTCA 1692
QY 1613 TTTGGAGCTGTGCTGCCCAACCACTTGAAGTCTTCAACTCTTCAAGATGAAGATGTCCG 1672
DB 1693 TTTGGAGCTGTGCTGCCCAACCACTTGAAGTCTTCAACTCTTCAAGATGAAGATGTCCG 1752
QY 1673 TCATCTTGGGGTGTGTGACATGCGCTTTTGGGGTGTGTCTTGGAGTCTTCAACCGAGTGC 1732
DB 1753 TCATCTTGGGGTGTGTGACATGCGCTTTTGGGGTGTGTCTTGGAGTCTTCAACCGAGTGC 1812
QY 1733 ACTTTGGGCGAGAGGACCGGCTGTGTGAGAGAGCGCTGCGGAGCTCACCTTCTGCTGG 1792
DB 1813 ACTTTGGGCGAGAGGACCGGCTGTGTGAGAGAGCGCTGCGGAGCTCACCTTCTGCTGG 1872


```
QY 953 TGAAGGCGCTGTACCTGGCCCTGAACACAGTGCAGCGTGAGCACCACGACCAAGTGCTCA 1012
Db 1033 TGAAGGCGCTGTACCTGGCCCTGAACACAGTGCAGCGTGAGCACCACGACCAAGTGCTCA 1092
QY 1013 TTGCGAGGCGCTGTGCTCTGTGCGAGACCTGCGCCGCTTGCAGAGGCGCTTGCAGGACA 1072
Db 1093 TTGCGAGGCGCTGTGCTCTGTGCGAGACCTGCGCCGCTTGCAGAGGCGCTTGCAGGACA 1152
QY 1073 GCTCGATGAGAGGAGTGAGTGCGGTGCTCAACCGATCCCTGCGCGGACATGCCGCC 1132
Db 1153 GCTCGATGAGAGGAGTGAGTGCGGTGCTCAACCGATCCCTGCGCGGACATGCCGCC 1212
QY 1133 CCACACTCATCGCACCAACCGCTTCAACGCGCAGCTTCCAGGSCATCGTGATCGTACG 1192
Db 1213 CCACACTCATCGCACCAACCGCTTCAACGCGCAGCTTCCAGGSCATCGTGATCGTACG 1272
QY 1193 GGTGGGCGCTACACGAGGCTCAACCCGCGCTCCCTACACCATCATCATCTTCCCTTCC 1252
Db 1273 GGTGGGCGCTACACGAGGCTCAACCCGCGCTCCCTACACCATCATCATCTTCCCTTCC 1332
QY 1253 TGTTGCTGTGATGTTGCGGAGATGCGGCGCACGCGCTGCTCATGTTCCTTCTTCCGCTGG 1312
Db 1333 TGTTGCTGTGATGTTGCGGAGATGCGGCGCACGCGCTGCTCATGTTCCTTCTTCCGCTGG 1392
QY 1313 CCATGCTCTTGGGAGAACGACCGGCTGTGAAGCGCGCAGAACGAGATCTGGCAGA 1372
Db 1393 CCATGCTCTTGGGAGAACGACCGGCTGTGAAGCGCGCAGAACGAGATCTGGCAGA 1452
QY 1373 CTTTCTTCAAGGCGCGCTACCTGCTCTCTCTTATGGGCGCTTCTTCCATCTACACCGCT 1432
Db 1453 CTTTCTTCAAGGCGCGCTACCTGCTCTCTCTTATGGGCGCTTCTTCCATCTACACCGCT 1512
QY 1433 TCATCTAACAGATGCTTCAGTGGCGCACACGAGATCTTCCCTCGCGGCTCGAGTGCG 1492
Db 1513 TCATCTAACAGATGCTTCAGTGGCGCACACGAGATCTTCCCTCGCGGCTCGAGTGCG 1572
QY 1493 CGGCATGSCCAACAGCTGTGCTGGAGTGATGCTTCTGCGCCAGACACAGTCTTGA 1552
Db 1573 CGGCATGSCCAACAGCTGTGCTGGAGTGATGCTTCTGCGCCAGACACAGTCTTGA 1632
QY 1553 CCTCGATCCCAAGCTACCGGTGTCTCTCGGACCTTACCCCTTTGGCATCGATCTTA 1612
Db 1633 CCTCGATCCCAAGCTACCGGTGTCTCTCGGACCTTACCCCTTTGGCATCGATCTTA 1692
QY 1613 TTTGGAGCTGCTGCGCAACCACTGAGCTTCTCACTCTTCAAGATGAAGATGTCG 1672
Db 1693 TTTGGAGCTGCTGCGCAACCACTGAGCTTCTCACTCTTCAAGATGAAGATGTCG 1752
QY 1673 TCATCTGGGCGCTGTGCACATGGCTTTGGGCTGTCTCGAGTCTTCAACCAAGTGC 1732
Db 1753 TCATCTGGGCGCTGTGCACATGGCTTTGGGCTGTCTCGAGTCTTCAACCAAGTGC 1812
QY 1733 ACTTTGGCAGAGGACCGGCTGCTGCTGGAGACGCTCGGAGCTCACTTCTGCTGG 1792
Db 1813 ACTTTGGCAGAGGACCGGCTGCTGCTGGAGACGCTCGGAGCTCACTTCTGCTGG 1872
QY 1793 GACTCTTGGTTACTCTGTTCTTAGTGCATCTCAAGTGGCTGTGTGCTGGGTGCCA 1852
Db 1873 GACTCTTGGTTACTCTGTTCTTAGTGCATCTCAAGTGGCTGTGTGCTGGGTGCCA 1932
QY 1853 GGGCGGCTCG--CCAGCATCTCATCTTCAATCAACATGTTCTCTTCTCCCA 1909
Db 1933 GGGCGGCTCGGGCCCGCAGCATCTCATCTTCAATCAACATGTTCTCTTCTCCCA 1992
QY 1910 GCCCAGCAACAGGCTGCTTACCCCGCAGAGGTGTCCAGGCACTGGTGGTCC 1969
Db 1993 GCCCAGCAACAGGCTGCTTACCCCGCAGAGGTGTCCAGGCACTGGTGGTCC 2052
QY 1970 TGGCCTTGGCATGTGCGCATCTGCTGCTTGGCACACCTTGGCATGCTGCTGCAACGCC 2029
Db 2053 TGGCCTTGGCCATGTGCGCATCTGCTGCTTGGCACACCTTGGCATGCTGCTGCAACGCC 2112
```

```
QY 2030 ACCGCGCGCGCTGCGGAGGAGGCGCGCTGACCGACAGAGGAGAAACAGGCGCGGTTCG 2089
Db 2113 ACCGCGCGCGCTGCGGAGGAGGCGCGCTGACCGACAGAGGAGAAACAGGCGCGGTTCG 2172
QY 2090 TGGACCTGCTGACGACATCTGTGAATGGTGGAGCTCCGATGAGAAAGGAGGAGGCGC 2149
Db 2173 TGGACCTGCTGACGACATCTGTGAATGGTGGAGCTCCGATGAGAAAGGAGGAGGCGC 2232
QY 2150 TGGATGATGAAGAGAGGCGGAGCTGCTGCCCTCCAGAGTGCTCATGACACAGGCGCATCC 2209
Db 2233 TGGATGATGAAGAGAGGCGGAGCTGCTGCCCTCCAGAGTGCTCATGACACAGGCGCATCC 2292
QY 2210 ACACCATCGAGTTCCTGCTGGGCTGCTCCAAACACGCGCTCTACCTGCGCGCTGTGG 2269
Db 2293 ACACCATCGAGTTCCTGCTGGGCTGCTCCAAACGCGCTCTACCTGCGCGCTGTGG 2352
QY 2270 CCTTGAGCCTTGGCCACGCGCAGCTGTCCGAGGTTCTGTGGGCCATGTGTGATGCGCATAG 2329
Db 2353 CCTTGAGCCTTGGCCACGCGCAGCTGTCCGAGGTTCTGTGGGCCATGTGTGATGCGCATAG 2412
QY 2330 GCTTGGGCTTGGGCGGAGGAGTGCGGCTGTGGTGTGTGCTGCCCATCTTTTGGCG 2389
Db 2413 GCTTGGGCTTGGGCGGAGGAGTGCGGCTGTGGTGTGTGCTGCCCATCTTTTGGCG 2472
QY 2390 CTTTGGCGGTGATGACCGTGGCTATCTGCTGTGATGAGAGGACTCTCAGCGCTTCTCTG 2449
Db 2473 CTTTGGCGGTGATGACCGTGGCTATCTGCTGTGATGAGAGGACTCTCAGCGCTTCTCTG 2532
QY 2450 AGCGCCTTGGCTGACCTGGGTGGAATTCAGAAACAAGTTCTACTCAGGACGCGGTACA 2509
Db 2533 AGCGCCTTGGCTGACCTGGGTGGAATTCAGAAACAAGTTCTACTCAGGACGCGGTACA 2592
QY 2510 AGCTGAGTCCCTTCACTTCCGCTGCGCACAGATGACTAGGCGCCACTGAGAGTCTTGGCAG 2569
Db 2593 AGCTGAGTCCCTTCACTTCCGCTGCGCACAGATGACTAGGCGCCACTGAGAGTCTTGGCAG 2652
QY 2570 ACCTCTTCTTCCGACCTCTGAGCAGGAGGAATAAGACGTCGCGCTGGCA 2623
Db 2653 ACCTCTTCTTCCGACCTCTGAGCAGGAGGAATAAGACGTCGCGCTGGCA 2706
```

RESULT 10

```
US-09-949-016-1330
; Sequence 1330, Application US/09949016
; Patent No.: 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1330
; LENGTH: 2457
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1330
```

Query Match 63.3%; Score 1670; DB 3; Length 2457;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2060; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 561 CTTTGTGAGGTGCGGTGAGGCCCAAGGCGCTCGCCCTAGAGCGCTCTCTGGAG 620

Db 392 CTTTGTGAGGTGCGGTGAGGCCCAAGGCGCTCGCCCTAGAGCGCTCTCTGGAG 451

FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1331
LENGTH: 2457
TYPE: DNA
ORGANISM: Human
US-09-949-016-1331

Query Match 63.3%; Score 1670; DB 3; Length 2457;

Best Local Similarity 99.7%; Pred. No. 0;
Matches 2060; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

| | | | |
|----|------|--|------|
| QY | 561 | CTTTGTGGAGGTGCGGTGGAGGCCCAAGGCCCTCGCCCTAGAGCGCTGCTGGAG | 620 |
| DB | 392 | CTTTGTGGAGGTGCGGTGGAGGCCCAAGGCCCTCGCCCTAGAGCGCTGCTGGAG | 451 |
| QY | 621 | GGCTGCGCGGCTTCTCATTTGCCAGCTTCAGGGAGCTGGAGCGCGCTGGAGCACCC | 680 |
| DB | 452 | GGCTGCGCGGCTTCTCATTTGCCAGCTTCAGGGAGCTGGAGCGCGCTGGAGCACCC | 511 |
| QY | 681 | CGTGACGGCGAGCCAGCCACGTCGATGACCTTCTCATCTCCTACTCGGGGTGAGCAGAT | 740 |
| DB | 512 | CGTGACGGCGAGCCAGCCACGTCGATGACCTTCTCATCTCCTACTCGGGGTGAGCAGAT | 571 |
| QY | 741 | CGGACGAAGATCCGCAAGATCACGGAATCTGTTTCCACTGCGCAGCTTCCCGTTTCTGCA | 800 |
| DB | 572 | CGGACGAAGATCCGCAAGATCACGGAATCTGTTTCCACTGCGCAGCTTCCCGTTTCTGCA | 631 |
| QY | 801 | GCAGGAGGAGCGCGCTCGGGCCCTGCAGCAGCTGCAACAGCAGCAGGAGCTGCA | 860 |
| DB | 632 | GCAGGAGGAGCGCGCGCTCGGGCCCTGCAGCAGCTGCAACAGCAGCAGGAGCTGCA | 691 |
| QY | 861 | GGAGGTCTCTCGGGGAGACAGAGCGGTTCTTGAGCCAGGTGCTAGGCCGGGTGCTGAGCT | 920 |
| DB | 692 | GGAGGTCTCTCGGGGAGACAGAGCGGTTCTTGAGCCAGGTGCTAGGCCGGGTGCTGAGCT | 751 |
| QY | 921 | GCTGCCGACAGGAGGTGAGTTCACAGATGAAAGGCGGTGATCTGGCCCTGAAACA | 980 |
| DB | 752 | GCTGCCGACAGGAGGTGAGTTCACAGATGAAAGGCGGTGATCTGGCCCTGAAACA | 811 |
| QY | 981 | GTGACGCTGAGCACACACCAAGTGCTCTATTGCCGAGGCTGGTGCTCTGTGCGAGA | 1040 |
| DB | 812 | GTGACGCTGAGCACACACCAAGTGCTCTATTGCCGAGGCTGGTGCTCTGTGCGAGA | 871 |
| QY | 1041 | CTTGCCCGCTGACAGAGGCGCTGCGGACAGCTCGATGGAGGGAGTGAAGTCCGT | 1100 |
| DB | 872 | CTTGCCCGCTGACAGAGGCGCTGCGGACAGCTCGATGGAGGGAGTGAAGTCCGT | 931 |
| QY | 1101 | GGCTACCGCATCCCTGCGGGACATGCCCCCACAATCTATCCGACCAACCGCTTAC | 1160 |
| DB | 932 | GGCTACCGCATCCCTGCGGGACATGCCCCCACAATCTATCCGACCAACCGCTTAC | 991 |
| QY | 1161 | GGCCAGCTTCCAGGGCATCGTGATCGCTAGGCGGTGGCGCTACACAGAGGTCAACCC | 1220 |
| DB | 992 | GGCCAGCTTCCAGGGCATCGTGATCGCTAGGCGGTGGCGCTACACAGAGGTCAACCC | 1051 |
| QY | 1221 | CGCTCCCTACACATCATCACTTCCCTTCTGTTTGTGTTGCTGTTGCTGGGAGTGGG | 1280 |
| DB | 1052 | CGCTCCCTACACATCATCACTTCCCTTCTGTTTGTGTTGCTGTTGCTGGGAGTGGG | 1111 |
| QY | 1281 | CCACGGGCTGCTATGTTCTTTCGCTTGGCCATGGTCTTGGGAGAACGACCGGC | 1340 |
| DB | 1112 | CCACGGGCTGCTATGTTCTTTCGCTTGGCCATGGTCTTGGGAGAACGACCGGC | 1171 |

| | | | |
|----|------|---|------|
| QY | 1341 | TGTGAAGCCGCGCAGAACGAGATCTGGCAGACTTTCTTACGGGCGCGCTACTGCTCCT | 1400 |
| DB | 1172 | TGTGAAGCCGCGCAGAACGAGATCTGGCAGACTTTCTTACGGGCGCGCTACTGCTCCT | 1231 |
| QY | 1401 | GCTTATGGGCTGTTCTCCATCTACACGGGCTTCTATACAAAGAGTGTCTTCACTGCGC | 1460 |
| DB | 1232 | GCTTATGGGCTGTTCTCCATCTACACGGGCTTCTATACAAAGAGTGTCTTCACTGCGC | 1291 |
| QY | 1461 | CACGAGCATTTTCCCTCGGGCTGAGTGTGGCGCATGGCCAAACAGTCTGGTGGAG | 1520 |
| DB | 1292 | CACGAGCATTTTCCCTCGGGCTGAGTGTGGCGCATGGCCAAACAGTCTGGTGGAG | 1351 |
| QY | 1521 | TGATGATTTCTGGGCGCAGCACACCATGCTTACCTGGATCCCAAGTCAACGCTGCTT | 1580 |
| DB | 1352 | TGATGATTTCTGGGCGCAGCACACCATGCTTACCTGGATCCCAAGTCAACGCTGCTT | 1411 |
| QY | 1581 | CCTGGGACCTTACCCCTTTGGCATCGATCTTATTTGGAGCCTGGCTGCCAAACACTTGAG | 1640 |
| DB | 1412 | CCTGGGACCTTACCCCTTTGGCATCGATCTTATTTGGAGCCTGGCTGCCAAACACTTGAG | 1471 |
| QY | 1641 | CTTCTCTCAACTCTCTTCAAGATGAAGTGTCCGTCACTCTGGGCGTGTGCAATGGCTT | 1700 |
| DB | 1472 | CTTCTCTCAACTCTCTTCAAGATGAAGTGTCCGTCACTCTGGGCGTGTGCAATGGCTT | 1531 |
| QY | 1701 | TGGGGTGTCTCGAGTCTTCAACACAGTGACCTTTGGCCAGAGGACCCGGCTGTGT | 1760 |
| DB | 1532 | TGGGGTGTCTCGAGTCTTCAACACAGTGACCTTTGGCCAGAGGACCCGGCTGTGT | 1591 |
| QY | 1761 | GGAGACGCTGCGGAGCTCACCTTCTGCTGGGACTCTTTCGGTTTACCTCGTGTCTAT | 1820 |
| DB | 1592 | GGAGACGCTGCGGAGCTCACCTTCTGCTGGGACTCTTTCGGTTTACCTCGTGTCTAT | 1651 |
| QY | 1821 | CATCTACAAGTGGCTGTGTGCTGGGCTGCGAGGCGCGCTCG---CCAGCATCTCAT | 1877 |
| DB | 1652 | CATCTACAAGTGGCTGTGTGCTGGGCTGCGAGGCGCGCTCG---CCAGCATCTCAT | 1711 |
| QY | 1878 | CCACTTTCATCAACATGTTCTCTTCTCCACAGCCCCAGCAACAGGCTGCTTACCCCCG | 1937 |
| DB | 1712 | CCACTTTCATCAACATGTTCTCTTCTCCACAGCCCCAGCAACAGGCTGCTTACCCCCG | 1771 |
| QY | 1938 | GCAGGAGGTGTCACAGGCCACGCTGCTGCTTCTGCGCTTTGGCCATGCTGCTCCTGCT | 1997 |
| DB | 1772 | GCAGGAGGTGTCACAGGCCACGCTGCTGCTTCTGCGCTTTGGCCATGCTGCTCCTGCT | 1831 |
| QY | 1998 | GCTTGGCACACCCCTGCACTGCTGCAACCGCCACCGCCGCTGCGAGAGGAGGCGCCG | 2057 |
| DB | 1832 | GCTTGGCACACCCCTGCACTGCTGCAACCGCCACCGCCGCTGCGAGAGGAGGCGCCG | 1891 |
| QY | 2058 | TGACCGACAGAGGAGAAACAGGCGCGGTTGCTGACCTGCTGACGCACTGTGTAATGG | 2117 |
| DB | 1892 | TGACCGACAGAGGAGAAACAGGCGCGGTTGCTGACCTGCTGACGCACTGTGTAATGG | 1951 |
| QY | 2118 | CTTGGAGCTCCGATGAGGAAAGGCAAGGCGCGGCTGATGATGAAGAGGAGGCGGAGCTCGT | 2177 |
| DB | 1952 | CTTGGAGCTCCGATGAGGAAAGGCAAGGCGCGGCTGATGATGAAGAGGAGGCGGAGCTCGT | 2011 |
| QY | 2178 | CCCTCTCGAGGCTGCTCATGCAACGACCATCCACACCATCGAGTTCCTGCTGGCTGCGT | 2237 |
| DB | 2012 | CCCTCTCGAGGCTGCTCATGCAACGACCATCCACACCATCGAGTTCCTGCTGGCTGCGT | 2071 |
| QY | 2238 | CTTCAACACCGCTCTTACCTCGGCTGTGGGCGCTTGAGCCTTGGCCACGCGCCAGCTGTC | 2297 |
| DB | 2072 | CTTCAACACCGCTCTTACCTCGGCTGTGGGCGCTTGAGCCTTGGCCACGCGCCAGCTGTC | 2131 |
| QY | 2298 | CGAGGTCTGTGGGCGATGCTGATCGCATAGGCTTGGGCTTGGGCGGGAAGTGGGCGT | 2357 |
| DB | 2132 | CGAGGTCTGTGTGGGCGATGCTGATCGCATAGGCTTGGGCGCTTGGGCGGGAAGTGGGCGT | 2191 |
| QY | 2358 | GGCGGCTGTGTGCTGCTTCCCATCTTTCGCGCTTTCGCGTATGACGCTGGCTATCCT | 2417 |
| DB | 2192 | GGCGGCTGTGTGCTGCTTCCCATCTTTCGCGCTTTCGCGTATGACGCTGGCTATCCT | 2251 |
| QY | 2418 | GCTGTGATGGAGGAGCTCTCAGCCTTCTGTCACGCCCTTCTGCGGCTGCACTGGGTGGAAT | 2477 |

Job time : 492 secs

```
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-83505

Query Match      8.4%; Score 221; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.4e-69;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 254 CTTTCTGCAGGAGGAGGTGCGCGGGCTGGGTGGTCTGCCCCCGCCAAAGGGGAGGC 313
Db 70 CTTTCTGCAGGAGGAGGTGCGCGGGCTGGGTGGTCTGCCCCCGCCAAAGGGGAGGC 129

QY 314 TGC CGGACACCCACCCCGGACCTGCTGGCATCCAGGAGGAGACGGAGCGCTGGCCC 373
Db 130 TGC CGGACACCCACCCCGGACCTGCTGGCATCCAGGAGGAGACGGAGCGCTGGCCC 189

QY 374 AGGAGCTGGGGATGTGCGGGCAACACAGAGGCCCTGCGGGCCAGCTGCACCACTGC 433
Db 190 AGGAGCTGGGGATGTGCGGGCAACACAGAGGCCCTGCGGGCCAGCTGCACCACTGC 249

QY 434 AGCTCCACGCGCGGTGCTACGCCAGGGCCATGAACCTCAG 474
Db 250 AGCTCCACGCGCGGTGCTACGCCAGGGCCATGAACCTCAG 290

RESULT 15
US-09-949-016-107576
; Sequence 107576, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J, Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 107576
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-107576

Query Match      8.4%; Score 221; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.4e-69;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 254 CTTTCTGCAGGAGGAGGTGCGCGGGCTGGGTGGTCTGCCCCCGCCAAAGGGGAGGC 313
Db 70 CTTTCTGCAGGAGGAGGTGCGCGGGCTGGGTGGTCTGCCCCCGCCAAAGGGGAGGC 129

QY 314 TGC CGGACACCCACCCCGGACCTGCTGGCATCCAGGAGGAGACGGAGCGCTGGCCC 373
Db 130 TGC CGGACACCCACCCCGGACCTGCTGGCATCCAGGAGGAGACGGAGCGCTGGCCC 189

QY 374 AGGAGCTGGGGATGTGCGGGCAACACAGAGGCCCTGCGGGCCAGCTGCACCACTGC 433
Db 190 AGGAGCTGGGGATGTGCGGGCAACACAGAGGCCCTGCGGGCCAGCTGCACCACTGC 249

QY 434 AGCTCCACGCGCGGTGCTACGCCAGGGCCATGAACCTCAG 474
Db 250 AGCTCCACGCGCGGTGCTACGCCAGGGCCATGAACCTCAG 290
```

Search completed: June 30, 2006, 09:22:32

181 GCTTCGGTGGCGCTTCAGAGACGCTTTGTGTTGATGTTTGGCGCTGTGAGAGCTG 240
241 GAGAGACCTTCACTTCCTGACGAGAGAGGTGCGCGGGCTGGGTGCTGCTGCCCCCG 300
241 GAGAGACCTTCACTTCCTGACGAGAGAGGTGCGCGGGCTGGGTGCTGCTGCCCCCG 300
301 CCAAGGGGAGGCTGCGGCAACCCCAACCCCGGACCTGCTGGCATCCAGAGAGAGCG 360
301 CCAAGGGGAGGCTGCGGCAACCCCAACCCCGGACCTGCTGGCATCCAGAGAGAGCG 360
361 GAGCGCTGGCCAGAGAGCTGCGGATGTGCGGGCAACAGCAGAGCCCTGCGGGCCAG 420
361 GAGCGCTGGCCAGAGAGCTGCGGATGTGCGGGCAACAGCAGAGCCCTGCGGGCCAG 420
421 CTGACACAGCTGAGCTCAACGCGCGCTGCTACGCGAGGGCACTGAACCTCAGCTGGCA 480
421 CTGACACAGCTGAGCTCAACGCGCGCTGCTACGCGAGGGCACTGAACCTCAGCTGGCA 480
481 GCGGCCACACAGATGGGGCCTCAGAGAGAGCGCCCTGCTCAGAGCCCGGGGGCG 540
481 GCGGCCACACAGATGGGGCCTCAGAGAGAGCGCCCTGCTCAGAGCCCGGGGGCG 540
541 CACAGGACCTGAGGGTCAACTTTGTGGAGGTGCGGTGGAGCCCAAGGCCCCCTGCC 600
541 CACAGGACCTGAGGGTCAACTTTGTGGAGGTGCGGTGGAGCCCAAGGCCCCCTGCC 600
601 CTAGAGCGCTGCTCTGAGAGGGCTGCGCGGGCTTCTCATTTGCCAGCTTCAGGAGCTG 660
601 CTAGAGCGCTGCTCTGAGAGGGCTGCGCGGGCTTCTCATTTGCCAGCTTCAGGAGCTG 660
661 GAGCAGCGCTGGAGCACCCGCTGAGCGGGGAGCGCCAGCTGATGACCTTCCTCATC 720
661 GAGCAGCGCTGGAGCACCCGCTGAGCGGGGAGCGCCAGCTGATGACCTTCCTCATC 720
721 TCTACTGCGGTGAGCAGATCGGACAGAGATCCGCAAGATCACGAGCTGCTTCACTGC 780
721 TCTACTGCGGTGAGCAGATCGGACAGAGATCCGCAAGATCACGAGCTGCTTCACTGC 780
781 CAGCTTTCCTCGTTCTGACGAGGAGGCGCGCTCGGGCCCTGACGAGCTGCAAA 840
781 CAGCTTTCCTCGTTCTGACGAGGAGGCGCGCTCGGGCCCTGACGAGCTGCAAA 840
841 CAGCAGAGCAGAGCTGACGAGGCTCCTCGGGGAGACAGAGGGTTCCTGAGCAGCTG 900
841 CAGCAGAGCAGAGCTGACGAGGCTCCTCGGGGAGACAGAGGGTTCCTGAGCAGCTG 900
901 CTAGGCGGGTGTGACGCTGTGCGCCAGGCGAGGTGACAGTCCACAAGATGAAGGC 960
901 CTAGGCGGGTGTGACGCTGTGCGCCAGGCGAGGTGACAGTCCACAAGATGAAGGC 960
961 GTGTACTGGCCCTGAAACAGTGCAGCGTGAGCAACAGCAAGTGCCTCATTTGCCGAG 1020
961 GTGTACTGGCCCTGAAACAGTGCAGCGTGAGCAACAGCAAGTGCCTCATTTGCCGAG 1020
1021 GCTGTGCTGTGCGAGACTGCGCGCTGCGAGGAGCCCTGCGGACAGCTCGATG 1080
1021 GCTGTGCTGTGCGAGACTGCGCGCTGCGAGGAGCCCTGCGGACAGCTCGATG 1080
1081 GAGGAGGAGTGAAGTGCCTGAGTCAACCGCATCCCTGCGGGACATGCCCCCACTC 1140
1081 GAGGAGGAGTGAAGTGCCTGAGTCAACCGCATCCCTGCGGGACATGCCCCCACTC 1140
1141 ATCCGACCAACCGCTTCAAGCGAGCTTCAGAGGATCGTGGATCGCTACGGGTGGGC 1200
1141 ATCCGACCAACCGCTTCAAGCGAGCTTCAGAGGATCGTGGATCGCTACGGGTGGGC 1200
1201 CGCTACAGAGGTCAACCCGCTCCCTACACATCATCACTTCCTTCCTGCTTTGCT 1260
1201 CGCTACAGAGGTCAACCCGCTCCCTACACATCATCACTTCCTTCCTGCTTTGCT 1260
1261 GTGATGTTCCGGGATGTGGGCCACGCGGCTGCTCATGTTCTTCCCTTGCCCTTGCCATG 1320

1261 GTGATGTTCCGGGATGTGGGCCACGCGGCTGCTCATGTTCTTTCGCCCTTGCCATGGTC 1320
1321 CTTGCGGAGAACCGACCGGCTGTGAAAGCGCGCAAGAACGAGATCTGGCAGACTTTCTTC 1380
1321 CTTGCGGAGAACCGACCGGCTGTGAAAGCGCGCAAGAACGAGATCTGGCAGACTTTCTTC 1380
1381 AGGGGCGCTACCTGCTCTGCTTATGGGCTGTGTTCTCCATCTACACGGCTTCATCTAC 1440
1381 AGGGGCGCTACCTGCTCTGCTTATGGGCTGTGTTCTCCATCTACACGGGCTTCATCTAC 1440
1441 AACAGTGTCTCAGTTCGGGCCACACAGCATCTTCCCTCGGGCTGAGTGTGGCCGCAATG 1500
1441 AACAGTGTCTCAGTTCGGGCCACACAGCATCTTCCCTCGGGCTGAGTGTGGCCGCAATG 1500
1501 GCCAACAGTCTGCTGCTGAGTGATGCAATTCCTGGGCCAGCACACGATGCTTACCTGGAT 1560
1501 GCCAACAGTCTGCTGCTGAGTGATGCAATTCCTGGGCCAGCACACGATGCTTACCTGGAT 1560
1561 CCCAACGTCACCGGTGCTTCTCTGGGACCCCTACCCCTTTGGCATCGATCTATTTGGAGC 1620
1561 CCCAACGTCACCGGTGCTTCTCTGGGACCCCTACCCCTTTGGCATCGATCTATTTGGAGC 1620
1621 CTGGCTGCGCAACCACTTGAAGTTCCTCAACTCTTCAAGATGAAGATGCTCCATCTCTG 1680
1621 CTGGCTGCGCAACCACTTGAAGTTCCTCAACTCTTCAAGATGAAGATGCTCCATCTCTG 1680
1681 GCGCTGCTGACATGGCCCTTTGGGGTGTCTCTCGAGTCTTCAACACAGTGCACTTTGGC 1740
1681 GCGCTGCTGACATGGCCCTTTGGGGTGTCTCTCGAGTCTTCAACACAGTGCACTTTGGC 1740
1741 CAGAGGACCGGCTGCTCTGGAGACGCTGCGGAGCTCACCTTCTGCTGGGACTCTTTC 1800
1741 CAGAGGACCGGCTGCTCTGGAGACGCTGCGGAGCTCACCTTCTGCTGGGACTCTTTC 1800
1801 GGTACCTCGTGTCTCTAGTCTATCAAGTGGCTGTGTCTGGGCTGCCAGGGCCGCC 1860
1801 GGTACCTCGTGTCTCTAGTCTATCAAGTGGCTGTGTCTGGGCTGCCAGGGCCGCC 1860
1861 TCGCCAGGATCTCTCATCCACTTCAATCAATGTTCTTCTTCTCCACAGGCCAGCAAC 1920
1861 TCGCCAGGATCTCTCATCCACTTCAATCAATGTTCTTCTTCTCCACAGGCCAGCAAC 1920
1921 AGGCTGCTTACCCCGGACGAGTGTCTCAGGCAACGCTGCTGGTGTCTTGGCTTTGGCC 1980
1921 AGGCTGCTTACCCCGGACGAGTGTCTCAGGCAACGCTGCTGGTGTCTTGGCTTTGGCC 1980
1981 ATGTGCCCACCTCTGCTTGGCACACCCCTGCACTGCTGCAACGCGCACCGCGCGCC 2040
1981 ATGTGCCCACCTCTGCTTGGCACACCCCTGCACTGCTGCAACGCGCACCGCGCGCC 2040
2041 CTGCGGAGAGGCGCGCTGACCGACAGAGGAAACAAAGGCGGGTGTGCTGACCTGCTT 2100
2041 CTGCGGAGAGGCGCGCTGACCGACAGAGGAAACAAAGGCGGGTGTGCTGACCTGCTT 2100
2101 GACGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGAGGAGGGGCTTGATGATGAA 2160
2101 GACGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGAGGAGGGGCTTGATGATGAA 2160
2161 GAGGAGGCGGAGCTGCTCCCTTCCAGGTGTCTGACACCGAGCCATCCACACCATCGAG 2220
2161 GAGGAGGCGGAGCTGCTCCCTTCCAGGTGTCTGACACCGAGCCATCCACACCATCGAG 2220
2221 TTCTGCTGGGTGCTGCTCMAACACCGCTCTCACTTGGCTGTGGCCCTTGAGCCCTG 2280
2221 TTCTGCTGGGTGCTGCTCMAACACCGCTCTCACTTGGCTGTGGCCCTTGAGCCCTG 2280
2281 GCCCAGCCAGCTGCTCGAGGTTCTGTGGCCATGTGTGATGCGCATAGGCCCTGGGCCCTG 2340
2281 GCCCAGCCAGCTGCTCGAGGTTCTGTGGCCATGTGTGATGCGCATAGGCCCTGGGCCCTG 2340
2341 GCGCGGAGGTGGCGTGGCGCTGTGGTGTGGTGTGCTTCCCATCTTTTGGCCCTTTCGCTG 2400
2341 GCGCGGAGGTGGCGTGGCGCTGTGGTGTGGTGTGCTTCCCATCTTTTGGCCCTTTCGCTG 2400

QY 2401 ATGACCGTGGCTATCTCTGCTGGTATGAGAGGACTCTCAGCCTTCTGACGCGCCCTCGG 2460
DB 2401 ATGACCGTGGCTATCTCTGCTGGTATGAGAGGACTCTCAGCCTTCTGACGCGCCCTCGG 2460
QY 2461 CTGCACTGGGTGGAATTCACAGAACAGTCTTACTCAGGACGGGCTACAAGCTGAGTCCC 2520
DB 2461 CTGCACTGGGTGGAATTCACAGAACAGTCTTACTCAGGACGGGCTACAAGCTGAGTCCC 2520
QY 2521 TTCACTTCTGCTGCCACAGATGACTAGGGCCCACTGCAAGTCTCTGCCAGACCTCTTCTCT 2580
DB 2521 TTCACTTCTGCTGCCACAGATGACTAGGGCCCACTGCAAGTCTCTGCCAGACCTCTTCTCT 2580
QY 2581 GACCTCTGAGCAGAGAGAGAAATAACAGCGTCCGCCCTGGCAAAAAA 2640
DB 2581 GACCTCTGAGCAGAGAGAGAAATAACAGCGTCCGCCCTGGCAAAAAA 2640

RESULT 2
US-09-962-436-278
; Sequence 278, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962.436
; PRIOR FILING DATE: 2001-09-25
; PRIOR FILING DATE: 2000-09-25
; PRIOR FILING DATE: 2000-09-25
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 278
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-278

Query Match 100.0%; Score 2640; DB 3; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCTGCGGACGCGGACGCCAGCAGCGAGGCGCGCGCAGCACACCCGGGACCATG 60
DB 1 CGCGCTGCGGACGCGGACGCCAGCAGCGAGGCGCGCGCAGCACACCCGGGACCATG 60
QY 61 GGCTCATGTTCCGGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCCCACAGCGGCT 120
DB 61 GGCTCATGTTCCGGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCCCACAGCGGCT 120
QY 121 GCCTACACCTGCTGAGTGGGCTGGGAGCTGGGCTCTGAGTTTCAGAGACCTCAAC 180
DB 121 GCCTACACCTGCTGAGTGGGCTGGGAGCTGGGCTCTGAGTTTCAGAGACCTCAAC 180
QY 181 GCCTCGGTGAGCGCTTTCAGAGACGCTTTGTGTGATGTTTGGCGCTGTGAGGAGCTG 240
DB 181 GCCTCGGTGAGCGCTTTCAGAGACGCTTTGTGTGATGTTTGGCGCTGTGAGGAGCTG 240
QY 241 GAGAAGACCTTCACTTCTGAGAGAGAGGTGCGGCGGCTGGGTGCTTCTGCCCCG 300
DB 241 GAGAAGACCTTCACTTCTGAGAGAGAGGTGCGGCGGCTGGGTGCTTCTGCCCCG 300
QY 301 CCAAGGGAGGCTGCGGACCCCAACCCCGGAGCTGCTGCGCATCCAGAGGAGAGC 360
DB 301 CCAAGGGAGGCTGCGGACCCCAACCCCGGAGCTGCTGCGCATCCAGAGGAGAGC 360
QY 361 GAGCGCTGCGCCAGGAGCTGCGGAGTGTGCGGGGCAACACAGAGGCGCTGCGGGCCAG 420
DB 361 GAGCGCTGCGCCAGGAGCTGCGGAGTGTGCGGGGCAACACAGAGGCGCTGCGGGCCAG 420

QY 421 CTGCAACAGCTGCAAGCTTCAAGCGCGCGCTGCTAGCGCAGGGCCATGAACCTCAGCTGGCA 480
DB 421 CTGCAACAGCTGCAAGCTTCAAGCGCGCGCTGCTAGCGCAGGGCCATGAACCTCAGCTGGCA 480
QY 481 GCGCCGCACACAGATGGGGCTCAGAGAGAGCGCCCTGCTTCCAGGCCCCCGGGGGCGG 540
DB 481 GCGCCGCACACAGATGGGGCTCAGAGAGAGCGCCCTGCTTCCAGGCCCCCGGGGGCGG 540
QY 541 CACCAAGACCTGAGGGTCAAATTTGTGGCAGGTGCGGTGAGGCCCAACAGGCCCTTGC 600
DB 541 CACCAAGACCTGAGGGTCAAATTTGTGGCAGGTGCGGTGAGGCCCAACAGGCCCTTGC 600
QY 601 CTAGAGCGCTGCTGAGAGGGCTGCGCGGGCTTCTCATTTGSCAGCTTTCAGGAGCTG 660
DB 601 CTAGAGCGCTGCTGAGAGGGCTGCGCGGGCTTCTCATTTGSCAGCTTTCAGGAGCTG 660
QY 661 GAGCAGCGCTGAGACACCCCGTGAAGCGGAGCAGCAGCGTGGATGACCTTCTCATC 720
DB 661 GAGCAGCGCTGAGACACCCCGTGAAGCGGAGCAGCAGCGTGGATGACCTTCTCATC 720
QY 721 TCCTACTGGGGTGAAGATCGACAGAAAGATCGCAAGATCAAGGACTGCTTCCATGTC 780
DB 721 TCCTACTGGGGTGAAGATCGACAGAAAGATCGCAAGATCAAGGACTGCTTCCATGTC 780
QY 781 CAGCTCTTCCGCTTCTGAGCAGGAGGAGCGCGCTCGGGCCCTGCGAGCAGCTGCA 840
DB 781 CAGCTCTTCCGCTTCTGAGCAGGAGGAGCGCGCTCGGGCCCTGCGAGCAGCTGCA 840
QY 841 CAGCAGGCGGAGCTGAGGAGGTCTCGGGGAGACAGAGCGGTTCCTGAGCAGGTG 900
DB 841 CAGCAGGCGGAGCTGAGGAGGTCTCGGGGAGACAGAGCGGTTCCTGAGCAGGTG 900
QY 901 CTAGGCGGGTGTGCAAGTGTCTGCGCCAGGCGAGGTGCGAGTCCAAGAGTGAAGGCC 960
DB 901 CTAGGCGGGTGTGCAAGTGTCTGCGCCAGGCGAGGTGCGAGTCCAAGAGTGAAGGCC 960
QY 961 GTGTACTGCGCTTGAACAGCTGAGCAGCAGCAGCAGCAGTGCCTCATTTGCCGAG 1020
DB 961 GTGTACTGCGCTTGAACAGCTGAGCAGCAGCAGCAGCAGTGCCTCATTTGCCGAG 1020
QY 1021 GCCTGTGCTCTGTCGAGACCTGCGCGCTGCGAGGAGCGCTGCGGGACAGCTCGATG 1080
DB 1021 GCCTGTGCTCTGTCGAGACCTGCGCGCTGCGAGGAGCGCTGCGGGACAGCTCGATG 1080
QY 1081 GAGGAGGAGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1081 GAGGAGGAGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1141 ATCCGACCAACCGCTTACGGCGAGCTTCCAGGGCATCGTGGATCGCTAGGGCTGGG 1200
DB 1141 ATCCGACCAACCGCTTACGGCGAGCTTCCAGGGCATCGTGGATCGCTAGGGCTGGG 1200
QY 1201 CGCTACAGGAGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCTGCTGTTGCT 1260
DB 1201 CGCTACAGGAGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCTGCTGTTGCT 1260
QY 1261 GTGATGTTGGGGATGTGGGCCACCGGCTGCTCATGTTCTTCTGCGCTTGGCCATGTC 1320
DB 1261 GTGATGTTGGGGATGTGGGCCACCGGCTGCTCATGTTCTTCTGCGCTTGGCCATGTC 1320
QY 1321 CTTGGGAGAACCGACCGGCTGTGAAAGCGCGCAGAACAGATCTGGCAGACTTCTTTC 1380
DB 1321 CTTGGGAGAACCGACCGGCTGTGAAAGCGCGCAGAACAGATCTGGCAGACTTCTTTC 1380
QY 1381 AGGGGCGCTTACCTGCTTCTGCTTATGGGCTGTTCCTTCCATCTACACCGCTTCTATC 1440
DB 1381 AGGGGCGCTTACCTGCTTCTGCTTATGGGCTGTTCCTTCCATCTACACCGCTTCTATC 1440
QY 1441 AACAGTGTCTTCACTGCGCCACAGCATCTTCCCTCTCGGGCTGGAGTGTGGCGCCATG 1500
DB 1441 AACAGTGTCTTCACTGCGCCACAGCATCTTCCCTCTCGGGCTGGAGTGTGGCGCCATG 1500
QY 1501 GCCAACAGCTCTGGCTGGAGTGTGATGCTTCTGGCCCGCAGCAGCATGCTTACCTGGAT 1560

1501 GCGCAACGAGTCTGGCTGGAGTGATGATTCCTGGCCAGCACAGATGCTTACCTCGAT 1560
1561 CCCAAGCTCACCAGTGTCTTCTGGGACCGTACCTCTTGGGATCGATCTATTGGAGC 1620
1561 CCCAAGCTCACCAGTGTCTTCTGGGACCGTACCTCTTGGGATCGATCTATTGGAGC 1620
1621 CTGGCTGCCAACACATTGAGCTTCCTCAACTCTTCAAGATGAAGATGCCGTCACTCG 1680
1621 CTGGCTGCCAACACATTGAGCTTCCTCAACTCTTCAAGATGAAGATGCCGTCACTCG 1680
1681 GCGCTCGTGCACATGCGCTTTGGGGTGGTCTCGGAGTCTTCAACACAGTCACTTTGGC 1740
1681 GCGCTCGTGCACATGCGCTTTGGGGTGGTCTCGGAGTCTTCAACACAGTCACTTTGGC 1740
1741 CAGAGGACCGGCTGTCTGGAGCGCTGCCGGAGCTCACCTTCTCTGGGATCTTTC 1800
1741 CAGAGGACCGGCTGTCTGGAGCGCTGCCGGAGCTCACCTTCTCTGGGATCTTTC 1800
1801 GGTACCTCGTGTCTAGTCACTACAAAGTGGCTGTGTCTGGGCTGCCAGGGCGGC 1860
1801 GGTACCTCGTGTCTAGTCACTACAAAGTGGCTGTGTCTGGGCTGCCAGGGCGGC 1860
1861 TGCCCCAGCATCTCATCACTTCAATCAATGTTCTCTTCTCCACAGCCCCAGCAAC 1920
1861 TGCCCCAGCATCTCATCACTTCAATCAATGTTCTCTTCTCCACAGCCCCAGCAAC 1920
1921 AGGCTGCTTACCCCGGAGGAGTGGTCCAGGACACGCTGGTGTCTTGGCTTTGGCC 1980
1921 AGGCTGCTTACCCCGGAGGAGTGGTCCAGGACACGCTGGTGTCTTGGCTTTGGCC 1980
1981 ATGGTCCCATCTGCTGTGGACACACCCCTGACCTGTGACCGCCACCGCCGCGC 2040
1981 ATGGTCCCATCTGCTGTGGACACACCCCTGACCTGTGACCGCCACCGCCGCGC 2040
2041 CTGCGAGGAGGCGCGCTGACCGACAGGAGGAAACAAGCGCGGGTGTGACCTGCT 2100
2041 CTGCGAGGAGGCGCGCTGACCGACAGGAGGAAACAAGCGCGGGTGTGACCTGCT 2100
2101 GAGCATCTGTGAATGGCTGGAGTCCGATGAGGAAAGGAGGCGGCTGTGATGATGA 2160
2101 GAGCATCTGTGAATGGCTGGAGTCCGATGAGGAAAGGAGGCGGCTGTGATGATGA 2160
2161 GAGGAGGCGGAGTCTGCTCCCTCCGAGTGTCTATGACACAGGCGCATCCACACATCGAG 2220
2161 GAGGAGGCGGAGTCTGCTCCCTCCGAGTGTCTATGACACAGGCGCATCCACACATCGAG 2220
2221 TTCTGCTGGCTGTGCTTCCAAACACCGCTCTCTACCTGCGCTGTGGGCTTGAAGCTG 2280
2221 TTCTGCTGGCTGTGCTTCCAAACACCGCTCTCTACCTGCGCTGTGGGCTTGAAGCTG 2280
2281 GCCCAGCCAGTGTCCGAGTGTCTGTGGGCAATGGTATGCGGATAGGCTTGGGCTG 2340
2281 GCCCAGCCAGTGTCCGAGTGTCTGTGGGCAATGGTATGCGGATAGGCTTGGGCTG 2340
2341 GCGCGGAGGAGTGGGCTGTGGCTGTGGTCTCCCATCTTTGCCCTTTGGCGG 2400
2341 GCGCGGAGGAGTGGGCTGTGGCTGTGGTCTCCCATCTTTGCCCTTTGGCGG 2400
2401 ATGACCGTGGCTATCTGCTGTGTGATGAGGAGTCTCAGCCTTCTGACGCTTCTGAC 2460
2401 ATGACCGTGGCTATCTGCTGTGTGATGAGGAGTCTCAGCCTTCTGACGCTTCTGAC 2460
2461 CTGCACTGGTGGATTTCCAGAACAGTTCCTACTCAGGACCGGCTACAGCTGATGCC 2520
2461 CTGCACTGGTGGATTTCCAGAACAGTTCCTACTCAGGACCGGCTACAGCTGATGCC 2520
2521 TTCACTTGGCTGCCACAGATGACTAGGCGCCACTGACAGTCTCTGACAGCTCTTCT 2580
2521 TTCACTTGGCTGCCACAGATGACTAGGCGCCACTGACAGTCTCTGACAGCTCTTCT 2580
2581 GACCTTGAGGCGAGAGGAAATAAGACGGTCCGCTTGGCAAAAAAAGGAGGAGGAGG 2640

Db 2581 GACCTCTGAGGCGAGGAGAAATAAGACGGTCCGCTTGGCAAAAAAAGGAGGAGG 2640
RESULT 3
US-09-880-107-3363
; Sequence 3363, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIORITY FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3363
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U45285
US-09-880-107-3363
Query Match 100.0%; Score 2640; DB 3; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGGCTGCGCGAACCGGCGAGCCAGCAGCGGAGCGCGCGGCGAGCACACCCCGGGACCATG 60
Db 1 CCGGCTGCGCGAACCGGCGAGCCAGCAGCGGAGCGCGCGGCGAGCACACCCCGGGACCATG 60
QY 61 GCGTCCATGTTCCGGAGCGAGGAGTGGCCCTGGTCCAGCTCTTTTGGCCACAGCGGT 120
Db 61 GCGTCCATGTTCCGGAGCGAGGAGTGGCCCTGGTCCAGCTCTTTTGGCCACAGCGGT 120
QY 121 GCCTACACCTCGTGAAGTGGCTGGCGAGCTGGGCTCTCGTGAGTTTACAGACCTCAAC 180
Db 121 GCCTACACCTCGTGAAGTGGCTGGCGAGCTGGGCTCTCGTGAGTTTACAGACCTCAAC 180
QY 181 GCGTGGTGGCGCTTCCAGAGACGCTTTGTGTGTGATGTTTGGCGCTGTGAGGAGCTG 240
Db 181 GCGTGGTGGCGCTTCCAGAGACGCTTTGTGTGTGATGTTTGGCGCTGTGAGGAGCTG 240
QY 241 GAGAGACCTTCACTTCTGCGAGGAGGAGTGGCGCGGCTGGGCTGGTCTGCCCCCG 300
Db 241 GAGAGACCTTCACTTCTGCGAGGAGGAGTGGCGCGGCTGGGCTGGTCTGCCCCCG 300
QY 301 CCAAGGGGAGGCTGCGGCGACCCCAACCCCGGAGCTGCTGCGCATCCAGAGGAGAGC 360
Db 301 CCAAGGGGAGGCTGCGGCGACCCCAACCCCGGAGCTGCTGCGCATCCAGAGGAGAGC 360
QY 361 GAGCGCTGGGCCAGGAGCTGCGGATGTCGCGGCAACCCAGAGGCTTGGCGGCGCAG 420
Db 361 GAGCGCTGGGCCAGGAGCTGCGGATGTCGCGGCAACCCAGAGGCTTGGCGGCGCAG 420
QY 421 CTGCAACGCTGACCTCCAGCCCGCTTACGCCAGGCGCATGAACTCAGCTGGCA 480
Db 421 CTGCAACGCTGACCTCCAGCCCGCTTACGCCAGGCGCATGAACTCAGCTGGCA 480
QY 481 GCGGCCCAACACAGATGGGCTTCCAGAGGAGCGCCCTGCTCCAGGCGCCCGGGGGCGG 540
Db 481 GCGGCCCAACACAGATGGGCTTCCAGAGGAGCGCCCTGCTCCAGGCGCCCGGGGGCGG 540
QY 541 CACAGGAGCTGAGGCTCACTTTGTGGCAGTGGCGGAGCCCAACAGGCGCCCTGCC 600
Db 541 CACAGGAGCTGAGGCTCACTTTGTGGCAGTGGCGGAGCCCAACAGGCGCCCTGCC 600


```
/ FILE REFERENCE: 689290-71
/ CURRENT APPLICATION NUMBER: US/09/968,007A
/ CURRENT FILING DATE: 2001-10-02
/ PRIOR APPLICATION NUMBER: US/60/237,172
/ PRIOR FILING DATE: 2000-10-02
/ PRIOR APPLICATION NUMBER: US/60/237,173
/ PRIOR FILING DATE: 2000-10-02
/ PRIOR APPLICATION NUMBER: US/60/237,278
/ PRIOR FILING DATE: 2000-10-02
/ PRIOR APPLICATION NUMBER: US/60/237,294
/ PRIOR FILING DATE: 2000-10-02
/ PRIOR APPLICATION NUMBER: US/60/237,295
/ PRIOR FILING DATE: 2000-10-02
/ PRIOR APPLICATION NUMBER: US/60/237,316
/ PRIOR FILING DATE: 2000-10-02
/ NUMBER OF SEQ ID NOS: 1001
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 212
/ TYPE: DNA
/ LENGTH: 2655
/ ORGANISM: Homo sapiens
/ US-09-968-007A-212

Query Match      100.0%; Score 2640; DB 3; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCTGCGGACCGGACGAGCGAGCGGCGGCGGACGACACCCCGGGACCATG 60
DB 1 CGGCTGCGGACCGGACGAGCGAGCGGCGGCGGACGACACCCCGGGACCATG 60

QY 61 GCCTCCATGTTCCGAGCGAGAGGTGGCCCTGGTCCAGCTCTTTCGCCACAGCGCT 120
DB 61 GCCTCCATGTTCCGAGCGAGAGGTGGCCCTGGTCCAGCTCTTTCGCCACAGCGCT 120

QY 121 GCCTACACTGCTGAGTGGTGGCGAGCTGGGCTCGTGGAGTTCAGAGACCTCAAC 180
DB 121 GCCTACACTGCTGAGTGGTGGCGAGCTGGGCTCGTGGAGTTCAGAGACCTCAAC 180

QY 181 GCCTGGTGAGCGCTTCCAGAGACGCTTTGGTGTGATGTTGGCGTGTGAGAGCTG 240
DB 181 GCCTGGTGAGCGCTTCCAGAGACGCTTTGGTGTGATGTTGGCGTGTGAGAGCTG 240

QY 241 GAGAAGACCTTCACTTCTCGAGAGAGGTGGCGGCGGCTGGTCTCTGCCCCCG 300
DB 241 GAGAAGACCTTCACTTCTCGAGAGAGGTGGCGGCGGCTGGTCTCTGCCCCCG 300

QY 301 CAAAAGGGAGGCTGCGGCAACCCCAACCCCGGACCTGCTGCGCATCCAGAGAGAGC 360
DB 301 CAAAAGGGAGGCTGCGGCAACCCCAACCCCGGACCTGCTGCGCATCCAGAGAGAGC 360

QY 361 GAGCGCTGGCCAGGAGCTGGGATGTCGGGGCAACCCAGCGGCTGGCGGCCAG 420
DB 361 GAGCGCTGGCCAGGAGCTGGGATGTCGGGGCAACCCAGCGGCTGGCGGCCAG 420

QY 421 CTGCACACAGCTCAGCTCCAGCCCGCTGTACGCGAGGGCCATGAACCTCAGCTGGCA 480
DB 421 CTGCACACAGCTCAGCTCCAGCCCGCTGTACGCGAGGGCCATGAACCTCAGCTGGCA 480

QY 481 GCGGCCCAACAGATGGGCTCAGAGAGGACGCCCCCTGCTCCAGGCCCCCGGGGGCGG 540
DB 481 GCGGCCCAACAGATGGGCTCAGAGAGGACGCCCCCTGCTCCAGGCCCCCGGGGGCGG 540

QY 541 CACAGGACCTGAGGCTCAACTTTGTGGCAGGTGGCGTGGAGCCCAAGAGCCCTGCC 600
DB 541 CACAGGACCTGAGGCTCAACTTTGTGGCAGGTGGCGTGGAGCCCAAGAGCCCTGCC 600

QY 601 CTAGAGCGCTCTCTGAGGGGCTGCGCGGCTTCTCATTTGCCAGCTTCAAGGAGCTG 660
DB 601 CTAGAGCGCTCTCTGAGGGGCTGCGCGGCTTCTCATTTGCCAGCTTCAAGGAGCTG 660

QY 661 GAGCAGCGCTGAGCACCCTGAGCGGCGAGCCAGCCAGTGAATGACCTTCTCATC 720
DB 661 GAGCAGCGCTGAGCACCCTGAGCGGCGAGCCAGCCAGTGAATGACCTTCTCATC 720
```

```
DB 661 GAGCAGCGCTGAGCACCCTGAGCGGCGAGCCAGCCAGTGAATGACCTTCTCATC 720
QY 721 TCCTACTGGGTGAGCAGATCGGACAGAAAGATCCGCAAGATCACGGACTGCTTCACTGC 780
DB 721 TCCTACTGGGTGAGCAGATCGGACAGAAAGATCCGCAAGATCACGGACTGCTTCACTGC 780
QY 781 CAGGTCTTCCCGTTTCTCAGCAGGAGAGGCGCCCTCGGGGCTTGCAGCAGCTGCAA 840
DB 781 CAGGTCTTCCCGTTTCTCAGCAGGAGAGGCGCCCTCGGGGCTTGCAGCAGCTGCAA 840
QY 841 CAGCAGAGCCAGGAGCTCCTCGGGAGAGACAGAGCGGTTCTCTGAGCCAGGTG 900
DB 841 CAGCAGAGCCAGGAGCTCCTCGGGAGAGACAGAGCGGTTCTCTGAGCCAGGTG 900
QY 901 CTAGGCGGGTGTCTGAGCTGCTCCGCGAGGAGGTGCAGGTCCACAGATGAAGGCC 960
DB 901 CTAGGCGGGTGTCTGAGCTGCTCCGCGAGGAGGTGCAGGTCCACAGATGAAGGCC 960
QY 961 GTGTACTTGGCCCTGAACAGGTGAGCAGCACACGCAAGTGTCTCATTTGCCGAG 1020
DB 961 GTGTACTTGGCCCTGAACAGGTGAGCAGCACACGCAAGTGTCTCATTTGCCGAG 1020
QY 1021 GCCTGCTCTGTGCGAGACCTGCCCTCTGAGGAGGCTTCGCGGACAGCTCGATG 1080
DB 1021 GCCTGCTCTGTGCGAGACCTGCCCTCTGAGGAGGCTTCGCGGACAGCTCGATG 1080
QY 1081 GAGGAGGAGTGAAGTGGCTCAGCGCATCCCTGCGGGACATGCCGCCACACTC 1140
DB 1081 GAGGAGGAGTGAAGTGGCTCAGCGCATCCCTGCGGGACATGCCGCCACACTC 1140
QY 1141 ATCCGCACAAACCGCTTCAAGCGGAGCTTCAGGGGATCGTGATCGCTACGGCGTGGGC 1200
DB 1141 ATCCGCACAAACCGCTTCAAGCGGAGCTTCAGGGGATCGTGATCGCTACGGCGTGGGC 1200
QY 1201 CGCTACAGGAGGTCAACCCCGCTCCATACCAATCATCATCTTCCCTTCTCTGTTGCT 1260
DB 1201 CGCTACAGGAGGTCAACCCCGCTCCATACCAATCATCATCTTCCCTTCTCTGTTGCT 1260
QY 1261 GTGATGTTGGGGATGTGGGCAACGGGCTGCTCATGTTCTTTCGCGCTGGCCATGCTC 1320
DB 1261 GTGATGTTGGGGATGTGGGCAACGGGCTGCTCATGTTCTTTCGCGCTGGCCATGCTC 1320
QY 1321 CTTGCGGAGAAACCGGCTGTAAAGCGCGAGAAACAGATCTGGCAGACTTTCTTTC 1380
DB 1321 CTTGCGGAGAAACCGGCTGTAAAGCGCGAGAAACAGATCTGGCAGACTTTCTTTC 1380
QY 1381 AGGGGCGCTACCTGCTCTGCTTATGAGGCTGTTCTCATCTACACGGGTTCTATCTAC 1440
DB 1381 AGGGGCGCTACCTGCTCTGCTTATGAGGCTGTTCTCATCTACACGGGTTCTATCTAC 1440
QY 1441 AACGAGTCTTCAAGTCGCGCCACAGCATCTTCCCTCGGGGTGAGGTGGCGGCCATG 1500
DB 1441 AACGAGTCTTCAAGTCGCGCCACAGCATCTTCCCTCGGGGTGAGGTGGCGGCCATG 1500
QY 1501 GCCAACAGTCTGGGTGAGTGATGATTCCTTGGGCCAGCAGCATGCTTACCTTGGAT 1560
DB 1501 GCCAACAGTCTGGGTGAGTGATGATTCCTTGGGCCAGCAGCATGCTTACCTTGGAT 1560
QY 1561 CCCAAAGTCAAGTGTCTTCTGAGGACCTTACCCCTTTGGGATCGATCTTATTTGGAGC 1620
DB 1561 CCCAAAGTCAAGTGTCTTCTGAGGACCTTACCCCTTTGGGATCGATCTTATTTGGAGC 1620
QY 1621 CTGGCTGCAACCACTTGAAGTCTTCAAGTGAAGTGTCCGTCATCTCTG 1680
DB 1621 CTGGCTGCAACCACTTGAAGTCTTCAAGTGAAGTGTCCGTCATCTCTG 1680
QY 1681 GGCGTGTGCATATGGCCTTTGGGTGCTCTCGAGTCTTCAACACAGTGCATTTGGC 1740
DB 1681 GGCGTGTGCATATGGCCTTTGGGTGCTCTCGAGTCTTCAACACAGTGCATTTGGC 1740
QY 1741 CAGAGGACCGGCTGCTGAGAGCGCTGCGGAGCTCACCTTCTGCTGGGACTCTTC 1800
DB 1741 CAGAGGACCGGCTGCTGAGAGCGCTGCGGAGCTCACCTTCTGCTGGGACTCTTC 1800
```



```
QY 1801 GGTACCTCGTTCCTAGTCACTACAGTGGCTGTGTGTCTGGGCTGCCAGGGCGGCC 1860
Db 1801 GGTACCTCGTTCCTAGTCACTACAGTGGCTGTGTGTCTGGGCTGCCAGGGCGGCC 1860
QY 1861 TCGCCAGCATCTCATCTCATCTCATCAACATGTTCTCTCTCTCCACAGCCCCAGCAAC 1920
Db 1861 TCGCCAGCATCTCATCTCATCTCATCAACATGTTCTCTCTCTCCACAGCCCCAGCAAC 1920
QY 1921 AGGCTGTCTACCCCGGCGAGAGGTGTCCAGGCCACGCTGTGTGTCTGGCTTGGCC 1980
Db 1921 AGGCTGTCTACCCCGGCGAGAGGTGTCCAGGCCACGCTGTGTGTCTGGCTTGGCC 1980
QY 1981 ATGTGTCCTATCTGCTGTGTGTCACACCTCTGACCTGTGTCACCGCCACCGCGCGC 2040
Db 1981 ATGTGTCCTATCTGCTGTGTGTCACACCTCTGACCTGTGTCACCGCCACCGCGCGC 2040
QY 2041 CTGGGAGGAGGCGCGCTGACCCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100
Db 2041 CTGGGAGGAGGCGCGCTGACCCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100
QY 2101 GAGCATCTGTGAATGCTGAGCTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
Db 2101 GAGCATCTGTGAATGCTGAGCTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
QY 2161 GAGGAGGCGAGCTGCTCCCTCCGAGTGTCTATGACACAGGAGGAGGAGGAGGAGGAGG 2220
Db 2161 GAGGAGGCGAGCTGCTCCCTCCGAGTGTCTATGACACAGGAGGAGGAGGAGGAGGAGG 2220
QY 2221 TTCTGCTGCTGCTGCTGCTCAACACCGCTCTTACCTGCGCTGTGGGCTGTGGGCTGT 2280
Db 2221 TTCTGCTGCTGCTGCTGCTCAACACCGCTCTTACCTGCGCTGTGGGCTGTGGGCTGT 2280
QY 2281 GCCACGCCAGCTGTCCGAGGTTCTGTGGGCCATGATGCGCATAGGCTGGGCTGT 2340
Db 2281 GCCACGCCAGCTGTCCGAGGTTCTGTGGGCCATGATGCGCATAGGCTGGGCTGT 2340
QY 2341 GGCGGAGGTGGGCTGGGCTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2400
Db 2341 GGCGGAGGTGGGCTGGGCTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2400
QY 2401 ATGACCTGCTATCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
Db 2401 ATGACCTGCTATCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
QY 2461 CTGACCTGGTGGAAATCCAGAACAGTTCCTACTCAGGCGGCTTACAGCTGAGTCCC 2520
Db 2461 CTGACCTGGTGGAAATCCAGAACAGTTCCTACTCAGGCGGCTTACAGCTGAGTCCC 2520
QY 2521 TTACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
Db 2521 TTACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2580
QY 2581 GACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
Db 2581 GACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
```

RESULT 5

US-10-305-720-1094

```
; Sequence 1094, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Sellhauer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 1094
```

```
; LENGTH: 2655
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GenBank ID No. US20040010136A1 g1245045
US-10-305-720-1094
```

```
Query Match 100.0%; Score 2640; DB 7; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CGCGCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
Db 1 CGCGCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 60
QY 61 GGCTCCATGTTCCGAGCGAGGAGGTGGCTCTGTCAGCTCTTCTTCTGCCCCACACGCGCT 120
Db 61 GGCTCCATGTTCCGAGCGAGGAGGTGGCTCTGTCAGCTCTTCTTCTGCCCCACACGCGCT 120
QY 121 GCCTACACCTGCTGAGTCTGCGCTGGGCGAGCTGGGCTCTGCTGGAGTTTCAGAGCTCAAC 180
Db 121 GCCTACACCTGCTGAGTCTGCGCTGGGCGAGCTGGGCTCTGCTGGAGTTTCAGAGCTCAAC 180
QY 181 GCCTCGTGTAGCGCTTCCAGAGACGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
Db 181 GCCTCGTGTAGCGCTTCCAGAGACGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
QY 241 GAGAAGACCTTCACTTCTGAGGAGGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 300
Db 241 GAGAAGACCTTCACTTCTGAGGAGGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 300
QY 301 CCAAGGAGGAGGCTGCGCGCACCCCGCGGACCTGCTGCGCATCCAGAGGAGGAGCG 360
Db 301 CCAAGGAGGAGGCTGCGCGCACCCCGCGGACCTGCTGCGCATCCAGAGGAGGAGCG 360
QY 361 GAGGCGCTGCGCCAGGAGCTGCGGAGTGTGGGGGCAACAGAGGCGCTGCGGGGCGCGAG 420
Db 361 GAGGCGCTGCGCCAGGAGCTGCGGAGTGTGGGGGCAACAGAGGCGCTGCGGGGCGCGAG 420
QY 421 CTCACACAGCTGAGAGCTCCACGCGCGCTGTACGCCAGGCGCATGACCTCAGCTGCGCA 480
Db 421 CTGACACAGCTGAGAGCTCCACGCGCGCTGTACGCCAGGCGCATGACCTCAGCTGCGCA 480
QY 481 GCCGCCCACACAGATGGGGCTCAGAGAGGACGCGCTGCTTCCAGGCGCCCGCGGGGCGCG 540
Db 481 GCCGCCCACACAGATGGGGCTCAGAGAGGACGCGCTGCTTCCAGGCGCCCGCGGGGCGCG 540
QY 541 CACAGGACCTGAGGCTCAACTTTGTGGCAGGTGCGGTGAGGCGCCACAGGCGCCCTGCCC 600
Db 541 CACAGGACCTGAGGCTCAACTTTGTGGCAGGTGCGGTGAGGCGCCACAGGCGCCCTGCCC 600
QY 601 CTAGAGCGCTGCTGAGGCGCTGCGCGGCTTCTCTCATTCGCGAGCTTTCAGGAGGCTG 660
Db 601 CTAGAGCGCTGCTGAGGCGCTGCGCGGCTTCTCTCATTCGCGAGCTTTCAGGAGGCTG 660
QY 661 GAGCAGCGCTGAGGACACCCCGGTGACGGGCGAGCGACCGACCGTGGATGACCTTCTCATC 720
Db 661 GAGCAGCGCTGAGGACACCCCGGTGACGGGCGAGCGACCGACCGTGGATGACCTTCTCATC 720
QY 721 TCCTACTGGGGTGTGAGAGATCGGACAGAGATCGCAAGATCAGGAGCTGCTTCCACTGC 780
Db 721 TCCTACTGGGGTGTGAGAGATCGGACAGAGATCGCAAGATCAGGAGCTGCTTCCACTGC 780
QY 781 CAGCTCTTCCGCTTCTGACAGAGGAGGCGCGCTCGGGGCGCTGCGAGCAGCTGCA 840
Db 781 CAGCTCTTCCGCTTCTGACAGAGGAGGAGGCGCGCTCGGGGCGCTGCGAGCAGCTGCA 840
QY 841 CAGCAGAGCAGGAGTGTGAGGAGGCTTCTCGGGGAGACAGAGGCTTCTTCTGAGCAGGT 900
Db 841 CAGCAGAGCAGGAGTGTGAGGAGGCTTCTCGGGGAGACAGAGGCTTCTTCTGAGCAGGT 900
QY 901 CTAGGCGGGTGTGTGAGCTGCTGCGCGCGGCGAGGCGAGGTGCGAGTCCACAAGATGAAGGCC 960
```


Db 901 CTAGGCGGGTGTCTGAGCTGCTCCGCGCAGGCGAGTGCAGTCCACAGATGAAGGCC 960
QY 961 GTGTACTTGGCCCTGAAACAGTGTGAGGTGAGCACAACGCAAAAGTGTCTCATTTGCCGAG 1020
Db 961 GTGTACTTGGCCCTGAAACAGTGTGAGGTGAGCACAACGCAAAAGTGTCTCATTTGCCGAG 1020
QY 1021 GCCTGTGTCTCTGCGAGACCTGCGCCCTGCGAGAGGCCCTGCGGACAGCTCGATG 1080
Db 1021 GCCTGTGTCTCTGCGAGACCTGCGCCCTGCGAGAGGCCCTGCGGACAGCTCGATG 1080
QY 1081 GAGGAGGAGTGTGAGTGTGCTACCGCATCCCTGCGGACATGCCCCCACAACCTC 1140
Db 1081 GAGGAGGAGTGTGAGTGTGCTACCGCATCCCTGCGGACATGCCCCCACAACCTC 1140
QY 1141 ATCCGACACCAACCGCTTACCGGCGAGCTTCCAGGGCATGTGTGATCGCTACGGCGTGGGC 1200
Db 1141 ATCCGACACCAACCGCTTACCGGCGAGCTTCCAGGGCATGTGTGATCGCTACGGCGTGGGC 1200
QY 1201 CGCTACAGAGGTCAACCCCGCTTACCGCATCATCACTTCCCTTCCCTTCTGTTTGT 1260
Db 1201 CGCTACAGAGGTCAACCCCGCTTACCGCATCATCACTTCCCTTCCCTTCTGTTTGT 1260
QY 1261 GTGATGTTCCGGGATGTGGCCACCGGCTGCTCATGTTCTTCTGCCCCCTGGCCATGTC 1320
Db 1261 GTGATGTTCCGGGATGTGGCCACCGGCTGCTCATGTTCTTCTTCCCTTCCCTTCCCTTCCCT 1320
QY 1321 CTTGGGAGAAACCGACCGGCTGTGAAAGCGCGCAGACGAGATCTTGGCAGACTTTCTTC 1380
Db 1321 CTTGGGAGAAACCGACCGGCTGTGAAAGCGCGCAGACGAGATCTTGGCAGACTTTCTTC 1380
QY 1381 AGGGGCGGCTTACCTGCTCTGCTTATGGGCGCTTCTTCCATCTACACCGGCTTCACTAC 1440
Db 1381 AGGGGCGGCTTACCTGCTCTGCTTATGGGCGCTTCTTCCATCTACACCGGCTTCACTAC 1440
QY 1441 AAGAGTGTCTCAGTGTGCGCCACGACATCTTCCCTCGGGCTGAGTGTGGCGGCATG 1500
Db 1441 AAGAGTGTCTCAGTGTGCGCCACGACATCTTCCCTCGGGCTGAGTGTGGCGGCATG 1500
QY 1501 GCCAACAGTCTGGCTGGAGTGTGATCTTCTGGGCGGCGCAGCACAGATGTCTTACCTGGAT 1560
Db 1501 GCCAACAGTCTGGCTGGAGTGTGATCTTCTGGGCGGCGCAGCACAGATGTCTTACCTGGAT 1560
QY 1561 CCACAGTCTCAGTGTCTTCTGGGACCTTCCCTTGGCATCGATCTTATTTGGAGC 1620
Db 1561 CCACAGTCTCAGTGTCTTCTGGGACCTTCCCTTGGCATCGATCTTATTTGGAGC 1620
QY 1621 CTGGCTGCCAACCACTTGTGCTTCTCAACTCTTCAAGATGAAGATGTCCGTCATCTCG 1680
Db 1621 CTGGCTGCCAACCACTTGTGCTTCTCAACTCTTCAAGATGAAGATGTCCGTCATCTCG 1680
QY 1681 GGGCTGTGCACATGCGCTTGGGCTGCTCGGAGTCTTCAACAGTGTCTTCAACAGTGTGCG 1740
Db 1681 GGGCTGTGCACATGCGCTTGGGCTGCTCGGAGTCTTCAACAGTGTCTTCAACAGTGTGCG 1740
QY 1741 CAGAGCACCGGCTGTGTGAGAGCGTCCGAGCTCACCTTCTGCTGGAGCTCTTC 1800
Db 1741 CAGAGCACCGGCTGTGTGAGAGCGTCCGAGCTCACCTTCTGCTGGAGCTCTTC 1800
QY 1801 GGTACCTGTGTCTTGTAGTCTCTCAAGTGTGTGTCTGGGCTGCCAGGGCGGCC 1860
Db 1801 GGTACCTGTGTCTTGTAGTCTCTCAAGTGTGTGTCTGGGCTGCCAGGGCGGCC 1860
QY 1861 TGCCCCAGATCTCATCACTTCTCATCAATGTTCTCTTCCACAGACGCCAGCAAC 1920
Db 1861 TGCCCCAGATCTCATCACTTCTCATCAATGTTCTCTTCCACAGACGCCAGCAAC 1920
QY 1921 AGGCTGTCTTACCCCCGAGAGGTGTCCAGGCCACGCTGTGTCTGTGCTTGGCTTGGCC 1980
Db 1921 AGGCTGTCTTACCCCCGAGAGGTGTCCAGGCCACGCTGTGTGTCTGTGCTTGGCTTGGCC 1980
QY 1981 ATGTGTGCCATCTGTGTGTGGCACACCCCTGCACTCTGTGACCGCCACCGCGCGC 2040

RESULT 6

US-10-641-643-916

; Sequence 916, Application US/10641643

; Publication No. US20040077003A1

; GENERAL INFORMATION:

; APPLICANT: Cocks, Benjamin G.

; Suan G. Stuart

; Jeffrey J. Seilhamer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL

; GENE EXPRESSION

; NUMBER OF SEQUENCES: 1508

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/641,643

; FILING DATE: 14-Aug-2003

; CLASSIFICATION: <Unknown>

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: <Unknown>

; FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 916:
SEQUENCE CHARACTERISTICS:
LENGTH: 2655 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 91245045
SEQUENCE DESCRIPTION: SEQ ID NO: 916 :
US-10-641-643-916

Query Match 100.0%; Score 2640; DB 8; Length 2655;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2640; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGGCGTCCGGGAGCGGAGCGAGCGAGGCGCGGCGAGCACACCGGGGAGCCATG 60
DB 1 CGGCGTCCGGGAGCGGAGCGGAGCGGAGGCGCGGCGAGCACACCGGGGAGCCATG 60
QY 61 GGCTCCATGTTCCGGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCACACAGCGGCT 120
DB 61 GGCTCCATGTTCCGGAGCGAGGAGGTGGCCCTGGTCCAGCTCTTTCTGCCACACAGCGGCT 120
QY 121 GCCTACACCTGGCTAGTCCGCTGGGCGAGCTGGGCGCTCGTGGAGTTCCAGACCTCAAC 180
DB 121 GCCTACACCTGGCTAGTCCGCTGGGCGAGCTGGGCGCTCGTGGAGTTCCAGACCTCAAC 180
QY 181 GCCTCGGTAGCGCTTCCAGAGACGCTTTGTGTGTTGATGTTGGCGCTGTGAGGAGCTG 240
DB 181 GCCTCGGTAGCGCTTCCAGAGACGCTTTGTGTGTTGATGTTGGCGCTGTGAGGAGCTG 240
QY 241 GAGAAGACCTTCACTTCTCCAGGAGGAGGTGCGGCGGCTGGGCTGGTCTTCCGCCCG 300
DB 241 GAGAAGACCTTCACTTCTCCAGGAGGAGGTGCGGCGGCTGGGCTGGTCTTCCGCCCG 300
QY 301 CCAAGGGGAGGCTCCGGGACACCCCAACCGCGGAGCTGCTGCGCATCCAGGAGGAGCG 360
DB 301 CCAAGGGGAGGCTCCGGGACACCCCAACCGCGGAGCTGCTGCGCATCCAGGAGGAGCG 360
QY 361 GAGCGCTGCCAGAGGCTGCGGAGTGCGGGGCAACAGCAGGCGCTGCGGGCCAG 420
DB 361 GAGCGCTGCCAGAGGCTGCGGAGTGCGGGGCAACAGCAGGCGCTGCGGGCCAG 420
QY 421 CTGCACAGCTGAGCTTCCAGCGCGCGCTGCTACGCCAGGCGCAATGAACCTCAGCTGGCA 480
DB 421 CTGCACAGCTGAGCTTCCAGCGCGCGCTGCTACGCCAGGCGCAATGAACCTCAGCTGGCA 480
QY 481 CGCGCCCAACAGATGGGCGCTCAGAGAGACGCGCTGCTCAGGCGCGCGGGGGCGG 540
DB 481 CGCGCCCAACAGATGGGCGCTCAGAGAGACGCGCTGCTCAGGCGCGCGGGGGCGG 540
QY 541 CACAGGACCTGAGGCTCAACTTTGTGGCAGGTGCGGTGGAGCCCAACAGGCGCTGCC 600
DB 541 CACAGGACCTGAGGCTCAACTTTGTGGCAGGTGCGGTGGAGCCCAACAGGCGCTGCC 600
QY 601 CTAGAGCGCTGCTCTGGAGGCGCTGCGCGGCTTCTCATTTGCGAGCTTTCAGGAGCTG 660
DB 601 CTAGAGCGCTGCTCTGGAGGCGCTGCGCGGCTTCTCATTTGCGAGCTTTCAGGAGCTG 660
QY 661 GAGCAGCGCTGAGACACCCCGTGAACCGGCGAGCAGCAGCAGTGGATGACCTTCTCATC 720
DB 661 GAGCAGCGCTGAGACACCCCGTGAACCGGCGAGCAGCAGCAGTGGATGACCTTCTCATC 720
QY 721 TCCTACTGGGCTGAGCAGATCGGACAGAGATCGCAAGATCATCGGAGTGTCTTCCATGTC 780
DB 721 TCCTACTGGGCTGAGCAGATCGGACAGAGATCGCAAGATCATCGGAGTGTCTTCCATGTC 780

DB 721 TCCTACTGGGCTGAGCAGATCGGACAGAGATCGCAAGATCATCGGAGTGTCTTCCATGTC 780
QY 781 CAGGTCTTCCCGTTTCTGACAGAGAGGAGGCGCGCTCGGGGCGCTGACAGAGTGCAG 840
DB 781 CAGGTCTTCCCGTTTCTGACAGAGAGGAGGCGCGCTCGGGGCGCTGACAGAGTGCAG 840
QY 841 CAGCAGAGCCAGGAGCTGACAGAGGTTCTCGGGAGACAGAGCGGTTCTTGAGCAGGTG 900
DB 841 CAGCAGAGCCAGGAGCTGACAGAGGTTCTCGGGAGAGGAGGTTCTTGAGCAGGTG 900
QY 901 CTAGGCGGCTGCTGACAGTCTGCGCCAGGCGAGGTGAGGTCCCAAGATGAAGGCC 960
DB 901 CTAGGCGGCTGCTGACAGTCTGCGCCAGGCGAGGTGAGGTCCCAAGATGAAGGCC 960
QY 961 GTGTACTGCGCCCTGAACCAAGTGCAGGTGAGCAACCAAGTGCCTCATTTGCCGAG 1020
DB 961 GTGTACTGCGCCCTGAACCAAGTGCAGGTGAGCAACCAAGTGCCTCATTTGCCGAG 1020
QY 1021 GCCTGGTGTCTGTGCGAGACCTGCGCGCTGACAGAGGCGCTGCGGAGCAGTCTGATG 1080
DB 1021 GCCTGGTGTCTGTGCGAGACCTGCGCGCTGACAGAGGCGCTGCGGAGCAGTCTGATG 1080
QY 1081 GAGGAGGAGTGCAGTGCCTGCGCTCACCGATCCCTGCGGGAGATGCCCCCAGACATC 1140
DB 1081 GAGGAGGAGTGCAGTGCCTGCGCTCACCGATCCCTGCGGGAGATGCCCCCAGACATC 1140
QY 1141 ATCCGACCAACCGCTTACGCGGCTTCCAGGCGCATCGTGGATGCTACGCGGTGGGC 1200
DB 1141 ATCCGACCAACCGCTTACGCGGCTTCCAGGCGCATCGTGGATGCTACGCGGTGGGC 1200
QY 1201 CGCTACAGAGGTCAACCCCGCTTCCACCATCATCATCTTCCCTTCTCTCTTCTTCT 1260
DB 1201 CGCTACAGAGGTCAACCCCGCTTCCACCATCATCATCTTCCCTTCTCTCTTCTTCT 1260
QY 1261 GTGATGTTCCGGGATGTGGGCGCATCGGCTGCTCATGTTCTTCTTCCGCGCTGGCCATGTC 1320
DB 1261 GTGATGTTCCGGGATGTGGGCGCATCGGCTGCTCATGTTCTTCTTCCGCGCTGGCCATGTC 1320
QY 1321 CTTTCGGAGAACCGCGCTGTGAAGCGCGCAGAACGAGATCTGGCAGACTTCTTCTTC 1380
DB 1321 CTTTCGGAGAACCGCGCTGTGAAGCGCGCAGAACGAGATCTGGCAGACTTCTTCTTC 1380
QY 1381 AGGGGCGCTACCTGCTCTGCTTATGGGCGCTTCTTCCATCTACACGGGCTTCTATCTAC 1440
DB 1381 AGGGGCGCTACCTGCTCTGCTTATGGGCGCTTCTTCCATCTACACGGGCTTCTATCTAC 1440
QY 1441 AACGAGTGTCTCAGTTCGCGCACACGATCTTCCCTCGGCGCTGGAGTGGCGGCGCATG 1500
DB 1441 AACGAGTGTCTCAGTTCGCGCACACGATCTTCCCTCGGCGCTGGAGTGGCGGCGCATG 1500
QY 1501 GCCAACCGAGTCTGGCTGGAGTGTGATGCAATCTTCCGCCAGCACACGATGCTTACCTGGAT 1560
DB 1501 GCCAACCGAGTCTGGCTGGAGTGTGATGCAATCTTCCGCCAGCACACGATGCTTACCTGGAT 1560
QY 1561 CCCAACGTCACCGGCTGCTTCTGGGACCTTACCCCTTTGGCATCGATCTTATTTGGAGC 1620
DB 1561 CCCAACGTCACCGGCTGCTTCTGGGACCTTACCCCTTTGGCATCGATCTTATTTGGAGC 1620
QY 1621 CTGGCTGCCAACCACTTGAAGTCTTCAAGTGAAGATGTCGTCATCTCTG 1680
DB 1621 CTGGCTGCCAACCACTTGAAGTCTTCAAGTGAAGATGTCGTCATCTCTG 1680
QY 1681 GGGCTGTGCACATGGCTTTTGGGGTGGTCTCTCGGAGTCTTCAACCAAGTGCATCTTGGC 1740
DB 1681 GGGCTGTGCACATGGCTTTTGGGGTGGTCTCTCGGAGTCTTCAACCAAGTGCATCTTGGC 1740
QY 1741 CAGAGGACCGGCTGCTGCTGGAGAGCGTCCCGGAGCTCACTTCTCTGCTGGAGCTCTTTC 1800
DB 1741 CAGAGGACCGGCTGCTGCTGGAGAGCGTCCCGGAGCTCACTTCTCTGCTGGAGCTCTTTC 1800
QY 1801 GGTTACTCTGTGTTCTTAGTCACTCAAGTGTGCTGTCTGGGCTGCCAGGCGCGCC 1860
DB 1801 GGTTACTCTGTGTTCTTAGTCACTCAAGTGTGCTGTCTGGGCTGCCAGGCGCGCC 1860


```
QY 841 CAGCAGAGCAGAGCTGACGAGGTCTCTCGGGAGACAGAGCGTCTCTGACGAGGTG 900
DB 841 CAGCAGAGCAGAGCTGACGAGGTCTCTCGGGAGACAGAGCGTCTCTGACGAGGTG 900
QY 901 CTAGGCGGGGTGTGAGCTGTGCTGCGCCAGAGGCAAGGTCCACAAGATGAAGGCC 960
DB 901 CTAGGCGGGGTGTGAGCTGTGCTGCGCCAGAGGCAAGGTCCACAAGATGAAGGCC 960
QY 961 GTGTACTCTGGCCCTGAACAGGTGACGCTGAGCACCACCAACAAGTGCCTCATTTGCCGAG 1020
DB 961 GTGTACTCTGGCCCTGAACAGGTGACGCTGAGCACCACCAACAAGTGCCTCATTTGCCGAG 1020
QY 1021 GCCTGTGTCTCTGTGCGAGACCTGCGCCGCTGACGAGGCGCTGCGGACAGCTCGATG 1080
DB 1021 GCCTGTGTCTCTGTGCGAGACCTGCGCCGCTGACGAGGCGCTGCGGACAGCTCGATG 1080
QY 1081 GAGGAGGAGTGTGAGTGTGCTGACCGCATCCCTGCGGAGCATGCCCCCACAATC 1140
DB 1081 GAGGAGGAGTGTGAGTGTGCTGACCGCATCCCTGCGGAGCATGCCCCCACAATC 1140
QY 1141 ATCCGACCAACCGCTTCAACCGGCTTCCAGGGCATCGTGGATCGCTACGGCGTGGG 1200
DB 1141 ATCCGACCAACCGCTTCAACCGGCTTCCAGGGCATCGTGGATCGCTACGGCGTGGG 1200
QY 1201 CGCTACAGAGGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCTCTGTTGCT 1260
DB 1201 CGCTACAGAGGTCAACCCCGCTCCCTACACCATCATCACTTCCCTTCTCTGTTGCT 1260
QY 1261 GTGATTTCCGGGATGTGGGCAACGGGCTGCTCAATGTTCTTCCCTGCGCCATGCTC 1320
DB 1261 GTGATTTCCGGGATGTGGGCAACGGGCTGCTCAATGTTCTTCCCTGCGCCATGCTC 1320
QY 1321 CTTGCGGAGAACCGCGCTGTGAAGCGCGGAGAGATCTGGCAGACTTCTTCTC 1380
DB 1321 CTTGCGGAGAACCGCGCTGTGAAGCGCGGAGAGATCTGGCAGACTTCTTCTC 1380
QY 1381 AGGGGCGGTACCTGCTCTGCTTATGGGCTGTTCTTCCATCTACACCGGCTTCTATC 1440
DB 1381 AGGGGCGGTACCTGCTCTGCTTATGGGCTGTTCTTCCATCTACACCGGCTTCTATC 1440
QY 1441 AACGAGTGTTCAGTGTGCGCCACAGCATCTTCCCTGCGGCTGGAGTGTGGCGCATG 1500
DB 1441 AACGAGTGTTCAGTGTGCGCCACAGCATCTTCCCTGCGGCTGGAGTGTGGCGCATG 1500
QY 1501 GCCAACGCTGTGGTGGAGTGTGATTCCTGGCCCGGACACACGATGCTTACCTGGAT 1560
DB 1501 GCCAACGCTGTGGTGGAGTGTGATTCCTGGCCCGGACACACGATGCTTACCTGGAT 1560
QY 1561 CCCAACGTCAACCGGTGTCTCTGGGACCTTACCCCTTTGGCATCGATCCTATTTGGAGC 1620
DB 1561 CCCAACGTCAACCGGTGTCTCTGGGACCTTACCCCTTTGGCATCGATCCTATTTGGAGC 1620
QY 1621 CTGGCTGCCAACCACTTGAAGTCTTCAACTCTTCAAGATGAAGTGTCCGTCATCTCTG 1680
DB 1621 CTGGCTGCCAACCACTTGAAGTCTTCAACTCTTCAAGATGAAGTGTCCGTCATCTCTG 1680
QY 1681 GCGCTGTGCACATGCGCTTTGGGGTGTCTCGAGTCTTCAACACGCTGCACTTTGGC 1740
DB 1681 GCGCTGTGCACATGCGCTTTGGGGTGTCTCGAGTCTTCAACACGCTGCACTTTGGC 1740
QY 1741 CAGAGCACCGGTGTCTGTGGAGCGCTGCGGAGCTCACCTTCTGCTGGGACTCTTTC 1800
DB 1741 CAGAGCACCGGTGTCTGTGGAGCGCTGCGGAGCTCACCTTCTGCTGGGACTCTTTC 1800
QY 1801 GGTTACTGTGTCTGTAGTCTATCTACAAGTGGCTGTGTGTGGCTGCGGAGCGCGCC 1860
DB 1801 GGTTACTGTGTGTCTGTAGTCTATCTACAAGTGGCTGTGTGTGGCTGCGGAGCGCGCC 1860
QY 1861 TCGCCAGCATCTCATCACTTCTCATCAATGTTCTCTTCTTCCACAGCCCAAGCAAC 1920
DB 1861 TCGCCAGCATCTCATCACTTCTCATCAATGTTCTCTTCTTCCACAGCCCAAGCAAC 1920
QY 1921 AGGCTGCTCTACCCCCCGGAGGAGTGGTCCAGGCCACGCTGGTGGTCTCTGGCC 1980
```

```
DB 1921 AGGCTGCTCTACCCCCCGGAGGAGTGGTCCAGGCCACGCTGGTGGTCTCTGGCC 1980
QY 1981 ATGGTGCCCATCTCTGCTGTGGGCAACCCCTGCACTCTGTGCAACGCCACCGCCGCGC 2040
DB 1981 ATGGTGCCCATCTCTGCTGTGGGCAACCCCTGCACTCTGTGCAACGCCACCGCCGCGC 2040
QY 2041 CTGCGGAGGAGGCGCGCTGACCGACAGAGGAGAAACAGGCGGCTGTGTGACCTGCT 2100
DB 2041 CTGCGGAGGAGGCGCGCTGACCGACAGAGGAGAAACAGGCGGCTGTGTGACCTGCT 2100
QY 2101 GACGCATCTGTGAATGGCTGGAGTCTCCATGAGGAGAAAGGAGGGGCTGTGATGATGA 2160
DB 2101 GACGCATCTGTGAATGGCTGGAGTCTCCATGAGGAGAAAGGAGGGGCTGTGATGATGA 2160
QY 2161 GAGGAGGCGGAGCTGCTCCCTCCGAGTGTCTATGCAACCGGCTTCCACACCATCGAG 2220
DB 2161 GAGGAGGCGGAGCTGCTCCCTCCGAGTGTCTATGCAACCGGCTTCCACACCATCGAG 2220
QY 2221 TTCTGCTGGGCTGCTCTCAACACCGCTTCTACCTGCGCTGTGGGCTGTGGGCTGTG 2280
DB 2221 TTCTGCTGGGCTGCTCTCAACACCGCTTCTACCTGCGCTGTGGGCTGTGGGCTGTG 2280
QY 2281 GCGCACGCGGCTGCTGCGAGTCTGTGGGCTATGTCGATGCGCATAGGCTTGGGCTG 2340
DB 2281 GCGCACGCGGCTGCTGCGAGTCTGTGGGCTATGTCGATGCGCATAGGCTTGGGCTG 2340
QY 2341 GCGCGGAGTGGGCTGCGGCTGTGTGTGTCCTGTCCTGTCCTGTCGCGCTTTGCGCTG 2400
DB 2341 GCGCGGAGTGGGCTGCGGCTGTGTGTGTCCTGTCCTGTCCTGTCGCGCTTTGCGCTG 2400
QY 2401 ATGACCGTGGCTATCTGCTGTGATGAGGAGCTCTCAGCCTTCTGACGCTTCTG 2460
DB 2401 ATGACCGTGGCTATCTGCTGTGATGAGGAGCTCTCAGCCTTCTGACGCTTCTG 2460
QY 2461 CTGCACTGGTGGAAATCCAGAACAAAGTTCTACTCAGGCACGGGCTACAAGCTGAGTCC 2520
DB 2461 CTGCACTGGTGGAAATCCAGAACAAAGTTCTACTCAGGCACGGGCTACAAGCTGAGTCC 2520
QY 2521 TTCACCTTGGTCCACAGATGCTAGGGCCCATGTCGAGGCTCTGCGAGACCTCTCTCT 2580
DB 2521 TTCACCTTGGTCCACAGATGCTAGGGCCCATGTCGAGGCTCTGCGAGACCTCTCTCT 2580
QY 2581 GACCTCTCAGGCGAGAGGAAATAAGACGCTTCCGCTTGGCAAAAAA 2640
DB 2581 GACCTCTCAGGCGAGAGGAAATAAGACGCTTCCGCTTGGCAAAAAA 2640
```

RESULT 10

US-10-145-012-1

; Sequence 1, Application US/10145012

; Publication No. US20030124614A1

; GENERAL INFORMATION:

; APPLICANT: UTU et al.

; TITLE OF INVENTION: T-CELL MEMBRANE PROTEIN (TIRC7), PEPTIDES AND ANTIBODIES DERIVED

; FILE REFERENCE: 4400-0105P

; CURRENT APPLICATION NUMBER: US/10/145,012

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 2488

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (537)..(2378)

US-10-145-012-1

Query Match 63.9%; Score 1687; DB 7; Length 2488;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2077; Conservative 0; Mismatches 3; Indels 3; Gaps 1;


```

; GENERAL INFORMATION:
; APPLICANT: UTKU, et al., NALAN
; TITLE OF INVENTION: METHODS OF TREATMENT USING AN ANTIBODY THAT SPECIFICALLY BINDS TO
; TITLE OF INVENTION: AN EXTRACELLULAR DOMAIN OF A TIRC7 MEMBRANE PROTEIN
; FILE REFERENCE: 1472/71099-ZA/JPW/AG
; CURRENT APPLICATION NUMBER: US/11/126,866
; CURRENT FILING DATE: 2005-05-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 2488
; TYPE: DNA
; ORGANISM: human
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (537) .. (2378)
US-11-126-866-1

```

```
Query Match      63.9%; Score 1687; DB 13; Length 2488;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2077; Conservative 0; Mismatches 3; Indels 3; Gaps 1;
```

| | | | |
|------|----|---|------|
| 561 | Qy | CTTTGTGGCAGGTGGCCGTGGAGCCCCACAAGAGCCCTTGCCCTAGAGCGCTGCTCTGGAG | 620 |
| 392 | Db | CTTTGTGGCAGGTGGCCGTGGAGCCCCACAAGAGCCCTTGCCCTAGAGCGCTGCTCTGGAG | 451 |
| 621 | Qy | GGCTGCGCGGGCTTCTCTATTCGACGTTCAGGGAGCTGGAGCAGCGCTGGAGCACCC | 680 |
| 452 | Db | GGCTGCGCGGGCTTCTCTATTCGACGTTCAGGGAGCTGGAGCAGCGCTGGAGCACCC | 511 |
| 681 | Qy | CGTGACGGGCGGACGACCGACGTGGATGACCTTCTCTCATCTCTTACCTGGGGTGAGCAGAT | 740 |
| 512 | Db | CGTGACGGGCGGACGACCGACGTGGATGACCTTCTCTCATCTCTTACCTGGGGTGAGCAGAT | 571 |
| 741 | Qy | CGGACAGAAGATCGCAAGATCAAGACTGCTTCCACTGCGCAGTCTTCGCCGTTTCTGCA | 800 |
| 572 | Db | CGGACAGAAGATCGCAAGATCAAGACTGCTTCCACTGCGCAGTCTTCGCCGTTTCTGCA | 631 |
| 801 | Qy | GCAGGAGGAGGCGCGCTTCGGGGCCCTGCGAGCAGCTGCTCAACAGCAGAGCCAGGAGCTGCA | 860 |
| 632 | Db | GCAGGAGGAGGCGCGCTTCGGGGCCCTGCGAGCAGCTGCTCAACAGCAGAGCCAGGAGCTGCA | 691 |
| 861 | Qy | GGAGGTCTCTCGGGGAGACAGAGCGGTTCTCTGAGCCAGGTGCTAGGCCGGGTGCTGCAGCT | 920 |
| 692 | Db | GGAGGTCTCTCGGGGAGACAGAGCGGTTCTCTGAGCCAGGTGCTAGGCCGGGTGCTGCAGCT | 751 |
| 921 | Qy | GCTGCCGCCAGGGCAGGTGTCAGGTTCACAAGATGAAGGCCGCTGTACTCTGCCCTGGAACCA | 980 |
| 752 | Db | GCTGCCGCCAGGGCAGGTGTCAGGTTCACAAGATGAAGGCCGCTGTACTCTGCCCTGGAACCA | 811 |
| 981 | Qy | GTGCAGGCTGAGCACACGACACAGTGCCTCATTTGCCAGGCTGTGTGCTCTGTGCGAGA | 1040 |
| 812 | Db | GTGCAGGCTGAGCACACGACACAGTGCCTCATTTGCCAGGCTGTGTGCTCTGTGCGAGA | 871 |
| 1041 | Qy | CCTGCCCGCCTTGACAGAGCCCTTGCGGACAGCTCGATCGAGGAGGAGTGAGTGCCGT | 1100 |
| 872 | Db | CCTGCCCGCCTTGACAGAGCCCTTGCGGACAGCTCGATCGAGGAGGAGTGAGTGCCGT | 931 |
| 1101 | Qy | GGCTCACCGCATCCCTTGCCGGGACATGCCCCCCACACTCATCGCACCAACCGCTTCAC | 1160 |
| 932 | Db | GGCTCACCGCATCCCTTGCCGGGACATGCCCCCCACACTCATCGCACCAACCGCTTCAC | 991 |
| 1161 | Qy | GGCCAGCTTCCAGGGCATCGTGATCGCTACGGCGTGGGCGGTACACAGAGGTCAACCC | 1220 |
| 992 | Db | GGCCAGCTTCCAGGGCATCGTGATCGCTACGGCGTGGGCGGTACACAGAGGTCAACCC | 1051 |
| 1221 | Qy | CGCTCCCTACACCATCATCACTTCCCTTCCCTGCTGCTGATGTTGCGGGATGTGGG | 1280 |
| 1052 | Db | CGCTCCCTACACCATCATCACTTCCCTTCCCTGCTGCTGATGTTGCGGGATGTGGG | 1111 |
| 1281 | Qy | CCACGGGCTGCTCAGTGTCTCTTCGCCCTTGGCCATGGTCTTTGCGGAGAACCGACCGC | 1340 |
| 1112 | Db | CCACGGGCTGCTCAGTGTCTCTTCGCCCTTGGCCATGGTCTTTGCGGAGAACCGACCGC | 1171 |

QY 2418 GCTGGTGAATGAGGAGCTCTCAGCCTTCTGCAAGCTGCTGGCTGCACTGGTGGAAAT 2477
Db |||||
2552 GCTGGTGAATGAGGAGCTCTCAGCCTTCTGCAAGCTGCTGGCTGCACTGGTGGAAAT 2311
QY |||||
2478 CCAGAACAAAGTTCTACTCAGGACAGGGCTACAAGCTGAGTCCCTTTACCTTCGCTGCCAC 2537
Db |||||
2312 CCAGAACAAAGTTCTACTCAGGACAGGGCTACAAGCTGAGTCCCTTTACCTTCGCTGCCAC 2371
QY |||||
2538 AGATGACTAGGGCCCACTGCAAGGTCTGCGCAGACCTCTTCTGACCTCTGAGGACAGGAG 2597
Db |||||
2372 AGATGACTAGGGCCCACTGCAAGGTCTGCGCAGACCTCTTCTGACCTCTGAGGACAGGAG 2431
QY |||||
2598 AGGAATAAGACGCTGCGCCCTGCAAAAAAAAAAAAAAAAAAAAA 2640
Db |||||
2432 AGGAATAAGACGCTGCGCCCTGCAAAAAAAAAAAAAAAAAAAAA 2474

RESULT 12

US-11-126-841A-1

; Sequence 1, Application US/11126841A

; Publication No. US20050271659A1

; GENERAL INFORMATION:

; APPLICANT: UTU, et al., NALAN

; TITLE OF INVENTION: METHODS OF TREATMENT USING AN ANTIBODY THAT SPECIFICALLY BINDS TO

; TITLE OF INVENTION: AN EXTRACELLULAR DOMAIN OF A TIRC7 MEMBRANE PROTEIN

; FILE REFERENCE: 1472/1099-ZB/JPW/AG

; CURRENT APPLICATION NUMBER: US/11/126,841A

; CURRENT FILING DATE: 2005-05-10

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 1

; LENGTH: 2488

; TYPE: DNA

; ORGANISM: human

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (537)..(2378)

US-11-126-841A-1

Query Match 63.9%; Score 1687; DB 15; Length 2488;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2077; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

QY 561 CTTTGTGGCAGGTGCGCTGAGGCCCAACAGGCCCTTGCCTAGAGCGCTGCTCTGGAG 620
Db |||||
392 CTTTGTGGCAGGTGCGCTGAGGCCCAACAGGCCCTTGCCTAGAGCGCTGCTCTGGAG 451
QY |||||
621 GGCCTGCGCGGCTTCTCTATTCAGCTTCAGGAGCTGGAGCGGCTGGAGCACCC 680
Db |||||
452 GGCCTGCGCGGCTTCTCTATTCAGCTTCAGGAGCTGGAGCGGCTGGAGCACCC 511
QY |||||
681 CGTGACGGGGCAGCAGCCACGCTGGATGACCTTCTCATCTCTTACTTGGGGTGAGCAGAT 740
Db |||||
512 CGTGACGGGGCAGCAGCCACGCTGGATGACCTTCTCATCTCTTACTTGGGGTGAGCAGAT 571
QY |||||
741 CGGACAGAAATCGGAAAGTACAGGACTGCTTCACTGCAAGCTTTCGGCTTTCGCA 800
Db |||||
572 CGGACAGAAATCGGAAAGTACAGGACTGCTTCACTGCAAGCTTTCGGCTTTCGCA 631
QY |||||
801 GCAGGAGGAGCGCCCTCGGGCCCTGAGCAGCTGCAAGCAGGAGCGGCTGGAGTCA 860
Db |||||
632 GCAGGAGGAGCGCCCTCGGGCCCTGAGCAGCTGCAAGCAGGAGCGGCTGGAGTCA 691
QY |||||
861 GGAGGTCTCTCGGGGAGACAGAGCGGTTCTTGAGCAGGTGCTAGGCCGGGTGCTGACGCT 920
Db |||||
692 GGAGGTCTCTCGGGGAGACAGAGCGGTTCTTGAGCAGGTGCTAGGCCGGGTGCTGACGCT 751
QY |||||
921 GCTGCGCCACAGGAGGTGAGGTCCACAGATGAAGCCGTGTAAGCTGCGCTTGAACCA 980
Db |||||
752 GCTGCGCCACAGGAGGTGAGGTCCACAGATGAAGCCGTGTAAGCTGCGCTTGAACCA 811
QY |||||
981 GTCCAGGTGAGCACCAAGTGCCTCATTTGCCAGGCTGCTGCTGCGAGA 1040
Db |||||

```
QY 2118 CTGGAGCTCCGATGAGAAAGGACAGGGCCCTGGATGATGAAGAGAGGCCGAGCTCGT 2177
Db 1952 CTGGAGCTCCGATGAGAAAGGACAGGGCCCTGGATGATGAAGAGAGGCCGAGCTCGT 2011
QY 2178 CCCTCCGAGGTGCTCATGCAACAGGCCATCCACACCATCGAGTTCTGCTCGGCTGCGT 2237
Db 2012 CCCTCCGAGGTGCTCATGCAACAGGCCATCCACACCATCGAGTTCTGCTCGGCTGCGT 2071
QY 2238 CTCCAAACCGGCTCTACTTGGGCTGTGGGCCCTGAGGCTGGGCCCAACGCCAGCTGTC 2297
Db 2072 CTCCAAACCGGCTCTACTTGGGCTGTGGGCCCTGAGGCTGGGCCCAACGCCAGCTGTC 2131
QY 2298 CGAGGTTCCTGTGGGCAATGATGCGATAGGCTTGGGCTTGGGCTGGGCGGAGGTGGGCT 2357
Db 2132 CGAGGTTCCTGTGGGCAATGATGCGATAGGCTTGGGCTTGGGCTGGGCGGAGGTGGGCT 2191
QY 2358 GCGGCTGTGGTGTGCTCCCATCTTTGCGGCTTTGCGGTGATGACCGTGGCTATCCT 2417
Db 2192 GCGGCTGTGGTGTGCTCCCATCTTTGCGGCTTTGCGGTGATGACCGTGGCTATCCT 2251
QY 2418 GCTGGTGAAGGAGACTCTAGCCTTCTGCAAGGCTGAGTCCCTTCCACCTTGGCTGCCAC 2477
Db 2252 GCTGGTGAAGGAGACTCTAGCCTTCTGCAAGGCTGAGTCCCTTCCACCTTGGCTGCCAC 2371
QY 2478 CAGAAACAAAGTTCTACTCAGGACAGGGCTACAGGCTGAGTCCCTTCCACCTTGGCTGCCAC 2537
Db 2312 CAGAAACAAAGTTCTACTCAGGACAGGGCTACAGGCTGAGTCCCTTCCACCTTGGCTGCCAC 2371
QY 2538 AGATGACTAGGCGCCACTGCAAGTCTGCGGACCTCTTCTGACCTCTGAGGCGAGGAG 2597
Db 2372 AGATGACTAGGCGCCACTGCAAGTCTGCGGACCTCTTCTGACCTCTGAGGCGAGGAG 2431
QY 2598 AGGATAAAGACGGTCCGCGCTGGCAAAAAAAAAAAAAAAAAAAAA 2640
Db 2432 AGGATAAAGACGGTCCGCGCTGGCAAAAAAAAAAAAAAAAAAAAA 2474

RESULT 13
US-10-145-012-12
; Sequence 12, Application US/10145012
; Publication No. US20030124614A1
; GENERAL INFORMATION:
; APPLICANT: UTOKU et al.
; TITLE OF INVENTION: T-CELL MEMBRANE PROTEIN (TIRC7), PEPTIDES AND ANTIBODIES DERIVED
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 4400-0105P
; CURRENT APPLICATION NUMBER: US/10/145,012
; CURRENT FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 2488
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (537) .. (2378)
US-10-145-012-12

Query Match 62.0%; Score 1636; DB 7; Length 2488;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2076; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 561 CTTTGTGGCAGGTGCGGTGGAGCCCAAGGCCCTTGCCCTAGAGCGCTGCTCTGGAG 620
Db 392 CTTTGTGGCAGGTGCGGTGGAGCCCAAGGCCCTTGCCCTAGAGCGCTGCTCTGGAG 451
QY 621 GGCCTCCGCGGCTTCTCATTTGCCAGCTTTCAGGAGCTGGAGCAGCGCTGGAGCACCC 680
Db 452 GGCCTCCGCGGCTTCTCATTTGCCAGCTTTCAGGAGCTGGAGCAGCGCTGGAGCACCC 511
QY 681 CGTGACGGGCGAGCCAGCCACGCGGTGATGACCTTCTCATCTCTACTGCGGTGAGCAGAT 740
```

```
Db 512 CQTGACGGCGAGCCAGCCACCTGGATGACCTTCTCTACTCTGGGTGAGCAGAT 571
QY 741 CGGACAGAAAGATCCGCAAGATCACGGACTGTCTTCCATCTGCCACGTTCTCCGTTTCTGCA 800
Db 572 CGGACAGAAAGATCCGCAAGATCACGGACTGTCTTCCATCTGCCACGTTCTCCGTTTCTGCA 631
QY 801 GCAGGAGGAGGCCCGCCCTCGGGGCTTCAGCAGCTGCAAGAGTGAAGGCCGTGTACTCTGGCCCTGAACCA 860
Db 632 GCAGGAGGAGGCCCGCCCTCGGGGCTTCAGCAGCTGCAAGAGTGAAGGCCGTGTACTCTGGCCCTGAACCA 691
QY 861 GGAGGTCTCTCGGGGAGACAGAGCGGTTCTTGAAGCCAGGTGCTAGGCCGGGTCTGTCAGCT 920
Db 692 GGAGGTCTCTCGGGGAGACAGAGCGGTTCTTGAAGCCAGGTGCTAGGCCGGGTCTGTCAGCT 751
QY 921 GCTGCGCCACAGGCGAGGTGTCAGGTTCACAAGATGAAGGCCGTGTACTCTGGCCCTGAACCA 980
Db 752 GCTGCGCCACAGGCGAGGTGTCAGGTTCACAAGATGAAGGCCGTGTACTCTGGCCCTGAACCA 811
QY 981 GTGCAAGCTGAGCACACAGCAAGTGCCTCATTTGCCGAGGCTTGGTGTCTGTGCGAGA 1040
Db 812 GTGCAAGCTGAGCACACAGCAAGTGCCTCATTTGCCGAGGCTTGGTGTCTGTGCGAGA 871
QY 1041 CCTGCCCGCCCTCGAGGAGGCCCTTGCAGGACAGCTCGATGAGGAGGAGTGAAGTCCCGT 1100
Db 872 CCTGCCCGCCCTCGAGGAGGCCCTTGCAGGACAGCTCGATGAGGAGGAGTGAAGTCCCGT 931
QY 1101 GAGTCACCGCATCCCTTGCAGGAGCATGCCCCCAGCACATCATCTCCGACCAACCGCTTCCAC 1160
Db 932 GAGTCACCGCATCCCTTGCAGGAGCATGCCCCCAGCACATCATCTCCGACCAACCGCTTCCAC 991
QY 1161 GAGTCACCGCATCCCTTGCAGGAGCATGCCCCCAGCACATCATCTCCGACCAACCGCTTCCAC 1220
Db 992 GAGTCACCGCATCCCTTGCAGGAGCATGCCCCCAGCACATCATCTCCGACCAACCGCTTCCAC 1051
QY 1221 CGCTCCCTACACCATCATCACCTTCCCTTCTGTTGCTGTGATGTTTCGGGAGATGTGG 1280
Db 1052 CGCTCCCTACACCATCATCACCTTCCCTTCTGTTGCTGTGATGTTTCGGGAGATGTGG 1111
QY 1281 CCAGGGGCTGCTCATGTTCTCTTTCGCTTGGCCATGCTTCTTGGGAGAACCGACCGGC 1340
Db 1112 CCAGGGGCTGCTCATGTTCTCTTTCGCTTGGCCATGCTTCTTGGGAGAACCGACCGGC 1171
QY 1341 TGTGAAGCCGCGCAGAACGAGATCTGGCAGACTTTCCTCAGGGGCGCTTACCTGCTCCT 1400
Db 1172 TGTGAAGCCGCGCAGAACGAGATCTGGCAGACTTTCCTCAGGGGCGCTTACCTGCTCCT 1231
QY 1401 GCTTATGGGCTGTTTCTCCATCTACACCGGCTTTCATCTACAAAGAGTGTCTTCAAGTCCGCG 1460
Db 1232 GCTTATGGGCTGTTTCTCCATCTACACCGGCTTTCATCTACAAAGAGTGTCTTCAAGTCCGCG 1291
QY 1461 CACACGATCTTCCCTCGGGCTGGAGTGTGGCCGATGGCCCAACAGTCTGGCTGGAG 1520
Db 1292 CACACGATCTTCCCTCGGGCTGGAGTGTGGCCGATGGCCCAACAGTCTGGCTGGAG 1351
QY 1521 TGATGATCTTCCCTGGGCCAGCACAGATGCTTACCTCTGGATCCCAAGCTCACCGGTGCTT 1580
Db 1352 TGATGATCTTCCCTGGGCCAGCACAGATGCTTACCTCTGGATCCCAAGCTCACCGGTGCTT 1411
QY 1581 CTTGGGACCTTACCCCTTTGGCATCGATCCTATTGAGGCTTGGCTGGTCCCAACCACTTGAG 1640
Db 1412 CTTGGGACCTTACCCCTTTGGCATCGATCCTATTGAGGCTTGGCTGGTCCCAACCACTTGAG 1471
QY 1641 CTTCTCACTCTTCAAGATGAAGTGTCCGTCTCTTGGGGCTGCTGCAATGAGCTT 1700
Db 1472 CTTCTCACTCTTCAAGATGAAGTGTCCGTCTCTTGGGGCTGCTGCAATGAGCTT 1531
QY 1701 TGGGGTGGTCTCGGAGTCTTCAACCAAGTGTCTTGGCCAGAGGACACCGGCTGTGCT 1760
Db 1532 TGGGGTGGTCTCGGAGTCTTCAACCAAGTGTCTTGGCCAGAGGACACCGGCTGTGCT 1591
QY 1761 GGAGACGCTGCGGAGCTCACCTTCTGTCTGGGACTTTCCTGCTTGGGACTTTCCTGCTGTCTAGT 1820
```

| | | |
|------|---|------|
| 1592 | GGAGACGCTCCGGAGCTCACCTTCTCTGCTGGGACCTCTTCGGTTACCTCTGTTGTTCTTAGT | 1651 |
| 1821 | CATCTTCAAGTGTGCTGTGTCTGTGGGTGCACAGGCGCGCTCG--CCACAGCATCTCTCAT | 1877 |
| 1652 | CATCTCAAGTGTGCTGTGTCTGTGGGTGCACAGGCGCGCTCTCGGCCCCACAGCATCTCTCAT | 1711 |
| 1878 | CCACTTTCATCAACATGTTCTCTTCTCCACAGCCCCAGCAACAGGCTGCTCTACCCCCG | 1937 |
| 1712 | CCACTTTCATCAACATGTTCTCTTCTCCACAGCCCCAGCAACAGGCTGCTCTACCCCCG | 1771 |
| 1938 | GCAGGAGTGGTCCAGGCCACGCTGTGTGTCTCTGGCTTGGCCATGTTGGTCCCATCTCTGTCT | 1997 |
| 1772 | GCAGGAGTGGTCCAGGCCACGCTGTGTGTCTCTGGCTTGGCCATGTTGGTCCCATCTCTGTCT | 1831 |
| 1998 | GCTTTGGCACACCCCTTGCACCTGTCTGCACCGCCACCGCGCGCGCTCTCGGAGGAGGCCCGC | 2057 |
| 1832 | GCTTTGGCACACCCCTTGCACCTGTCTGCACCGCCACCGCGCGCGCTCTCGGAGGAGGCCCGC | 1891 |
| 2058 | TGACCGACAGGAGGAAAAAAGGCGGGTTGCTGGACCTTCGCTGACGCATCTGTGAATGG | 2117 |
| 1892 | TGACCGACAGGAGGAAAAAAGGCGGGTTGCTGGACCTTCGCTGACGCATCTGTGAATGG | 1951 |
| 2118 | CTGGAGCTCCGATGAGGAAAAAGGACAGGGGCCCTTGGATGATGAAGAGGAGGCCGAGCTCGT | 2177 |
| 1952 | CTGGAGCTCCGATGAGGAAAAAGGACAGGGGCCCTTGGATGATGAAGAGGAGGCCGAGCTCGT | 2011 |
| 2178 | CCCTTCGGAGTGTCTCATGACACAGGCCATTCACACCATCGAGTTCTGCCCTGGGCTGCGT | 2237 |
| 2012 | CCCTTCGGAGTGTCTCATGACACAGGCCATTCACACCATCGAGTTCTGCCCTGGGCTGCGT | 2071 |
| 2238 | CTCCAACACCGCTCTTACCTGCGCCTGTGGGCGCTTGAGCCTGCGCCACGCCCAGAGTGTC | 2297 |
| 2072 | CTCCAACACCGCTCTTACCTGCGCCTGTGGGCGCTTGAGCCTGCGCCACGCCCAGAGTGTC | 2131 |
| 2298 | CGAGGTTCTGTGGGCCCATGTGTGATGCGCATAGGCTTGGGCTTGGGCGGGAGGTGGGCGT | 2357 |
| 2132 | CGAGGTTCTGTGGGCCCATGTGTGATGCGCATAGGCTTGGGCTTGGGCGGGAGGTGGGCGT | 2191 |
| 2358 | GGCGGCTGTGTGCTCGTCCCATCTTTTGGCGGCTTGTCCGCTGATGACCGTGCGTATCCCT | 2417 |
| 2192 | GGCGGCTGTGTGCTCGTCCCATCTTTTGGCGGCTTGTCCGCTGATGACCGTGCGTATCCCT | 2251 |
| 2418 | GCTGTGTATGAGGGGACTCTCAGCCTTCTCTGCACGCCCTGCGGCTGCACCTGGGTGGAAAT | 2477 |
| 2252 | GCTGTGTATGAGGGGACTCTCAGCCTTCTCTGCACGCCCTGCGGCTGCACCTGGGTGGAAAT | 2311 |
| 2478 | CCAGAACAAAGTTCTACTCAGGCACGGGCTACAAGCTGAGTTCCTTACCTTCGCTGCCAC | 2537 |
| 2312 | CCAGAACAAAGTTCTACTCAGGCACGGGCTACAAGCTGAGTTCCTTACCTTCGCTGCCAC | 2371 |
| 2538 | AGATGACTAGGGGCCACTTGAGGTCTCTGCCAGACCTCTTCTCTGACCTCTTGAGGCAAGGAG | 2597 |
| 2372 | AGATGACTAGGGGCCACTTGAGGTCTCTGCCAGACCTCTTCTCTGACCTCTTGAGGCAAGGAG | 2431 |
| 2598 | AGGAATAAGACGGTCCGCCCTGGCAAAAAAAAAAAAAAAAAAAAAA | 2640 |
| 2432 | AGGAATAAGACGGTCCGCCCTGGCAAAAAAAAAAAAAAAAAAAAAA | 2474 |

RESULT 14
 US-11-126-866-12
 ; Sequence 12, Application US/11126866
 ; Publication No. US20050220789A1
 ; GENERAL INFORMATION:
 ; APPLICANT: UTKU, et al., NALAN
 ; TITLE OF INVENTION: METHODS OF TREATMENT USING AN ANTIBODY THAT SPECIFICALLY BINDS TO
 ; TITLE OF INVENTION: AN EXTRACELLULAR DOMAIN OF A TIRC7 MEMBRANE PROTEIN
 ; FILE REFERENCE: 1472/71099-ZA/3PW/AG
 ; CURRENT APPLICATION NUMBER: US/11/126,866
 ; CURRENT FILING DATE: 2005-05-10
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 12

```
QY 1461 CACGACATCTTCCCTCGGCTGGAGTGTGGCGGCATGGCCAAACAGTCTGGCTGGAG 1520
Db 1292 CACGACATCTTCCCTCGGCTGGAGTGTGGCGGCATGGCCAAACAGTCTGGCTGGAG 1351
QY 1521 TGATGATTCCTGGGCCAGACACGATGCTTACCTTGGATCCCAACAGTCAACCGGTGTCTT 1580
Db 1352 TGATGATTCCTGGGCCAGACACGATGCTTACCTTGGATCCCAACAGTCAACCGGTGTCTT 1411
QY 1581 CTGGGACCTTACCTTTGGGATCGATTCATTTGGAGCTTGGCTGGCCAAACACATTGAG 1640
Db 1412 CTGGGACCTTACCTTTGGGATCGATTCATTTGGAGCTTGGCTGGCCAAACACATTGAG 1471
QY 1641 CTTCTCAACTCTTCAAGATGAAGATGTCGGTCATCTCGGGGCTGTCACATGACCTT 1700
Db 1472 CTTCTCAACTCTTCAAGATGAAGATGTCGGTCATCTCGGGGCTGTCACATGACCTT 1531
QY 1701 TGGGGTGGTCTCGGAGTCTTCAACACAGTGCATTTGGCCAGAGGACCGGCTGTGCT 1760
Db 1532 TGGGGTGGTCTCGGAGTCTTCAACACAGTGCATTTGGCCAGAGGACCGGCTGTGCT 1591
QY 1761 GAGAGCGTCCGGAGCTCACCTTCTGCTGGGACTCTTTCGGTTACTCTGTGTTCTAGT 1820
Db 1592 GAGAGCGTCCGGAGCTCACCTTCTGCTGGGACTCTTTCGGTTACTCTGTGTTCTAGT 1651
QY 1821 CATCTACAAGTGGCTGTGTGCTGGGCTGCGAGGGCGGCTCG---CCAGCATCCTCAT 1877
Db 1652 CATCTACAAGTGGCTGTGTGCTGGGCTGCGAGGGCGGCTCGGCCCCCAGCATCCTCAT 1711
QY 1878 CCACTTCATCAATGTTCTCTTCTCCACAGCCCCAGCAACAGGCTGCTTACCCCCG 1937
Db 1712 CCACTTCATCAATGTTCTCTTCTCCACAGCCCCAGCAACAGGCTGCTTACCCCCG 1771
QY 1938 GAGAGGTGTTCAGGCCACGCTGTGTGCTGTGCTTGGCCATGGTCCCATCCTGT 1997
Db 1772 GAGAGGTGTTCAGGCCACGCTGTGTGCTGTGCTTGGCCATGGTCCCATCCTGT 1831
QY 1998 GCTTGGCACACCCCTGCATCTGTCACCGCACCGCGCGCTGCGAGGAGGCGCCG 2057
Db 1832 GCTTGGCACACCCCTGCATCTGTCACCGCACCGCGCGCTGCGAGGAGGCGCCG 1891
QY 2058 TGACGACAGGAGGAAAAAAGGCGGGTGTGTGACCTGCTGACGATCTGTGAATGG 2117
Db 1892 TGACGACAGGAGGAAAAAAGGCGGGTGTGTGACCTGCTGACGATCTGTGAATGG 1951
QY 2118 CTGGAGCTCCGATGAGGAAAAAGCAGGGGCTTGGATGATGAAGAGAGGCGCAGCTCGT 2177
Db 1952 CTGGAGCTCCGATGAGGAAAAAGCAGGGGCTTGGATGATGAAGAGAGGCGCAGCTCGT 2011
QY 2178 CCCCTCCGAGGTGCTCATGCACAGGCCATCCACACCATCGAGTTCTGGCTGGGCTGGCT 2237
Db 2012 CCCCTCCGAGGTGCTCATGCACAGGCCATCCACACCATCGAGTTCTGGCTGGGCTGGCT 2071
QY 2238 CTCCAACACCGCTCTTACCTGGCTGTGGGCTTGGAGCTTGGCCACGCCAGCTGTC 2297
Db 2072 CTCCAACACCGCTCTTACCTGGCTGTGGGCTTGGAGCTTGGCCACGCCAGCTGTC 2131
QY 2298 CAGAGTTCTGTGGGCCATGATGCGCATAGGCTTGGGCTTGGGCGGAGAGTGGGCGT 2357
Db 2132 CAGAGTTCTGTGGGCCATGATGCGCATAGGCTTGGGCTTGGGCGGAGAGTGGGCGT 2191
QY 2358 GCGGCTGTGTGCTGGTCCCATCTTTGCGGCTTTTGGCGTGATGACCGTGGCTATCCT 2417
Db 2192 GCGGCTGTGTGCTGGTCCCATCTTTGCGGCTTTTGGCGTGATGACCGTGGCTATCCT 2251
QY 2418 GCTGTGATGAGGAGGACTCTCAGCCTTCTGACGCTTCTGACGCTTGGCTGCACTGGGTGGAAT 2477
Db 2252 GCTGTGATGAGGAGGACTCTCAGCCTTCTGACGCTTCTGACGCTTGGCTGCACTGGGTGGAAT 2311
QY 2478 CAGAACAAATTTCTACTCAGGACCGGGCTACAAGCTGAGTCCCTTCACTTCTGCTGCCAC 2537
Db 2312 CAGAACAAATTTCTACTCAGGACCGGGCTACAAGCTGAGTCCCTTCACTTCTGCTGCCAC 2371
```

```
QY 2538 AGATGACTAGGCGCCACTGACGCTCTGCAGACCTCTTCTTGACCTCTGAGGCAAGAG 2597
Db 2372 AGATGACTAGGCGCCACTGACGCTCTGCAGACCTCTTCTTGACCTCTGAGGCAAGAG 2431
QY 2598 AGGAATAAAGACGGTCCGCCCTGGCAAAAAAAGGAAAAA 2640
Db 2432 AGGAATAAAGACGGTCCGCCCTGGCAAAAAAAGGAAAAA 2474
```

RESULT 15

```
US-11-126-841A-12
; Sequence 12, Application US(11126841A
; Publication No. US20050271659A1
; GENERAL INFORMATION:
; APPLICANT: UTNU, et al., NALAN
; TITLE OF INVENTION: METHODS OF TREATMENT USING AN ANTIBODY THAT SPECIFICALLY BINDS 1
; FILE REFERENCE: AN EXTRACELLULAR DOMAIN OF A TIRC7 MEMBRANE PROTEIN
; FILE REFERENCE: 1472/71099-ZB/JPW/AG
; CURRENT APPLICATION NUMBER: US/11/126,841A
; CURRENT FILING DATE: 2005-05-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 12
; LENGTH: 2488
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic oligonucleotide
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (537)..(2378)
US-11-126-841A-12
```

Query Match 62.0%; Score 1636; DB 15; Length 2488;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2076; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

```
QY 561 CTTTGTGGAGTGTGCGTGGAGCCCAAGGCGCTGAGAGCGCTGCTCTGGAG 620
Db 392 CTTTGTGGAGTGTGCGTGGAGCCCAAGGCGCTGAGAGCGCTGCTCTGGAG 451
QY 621 GGCCTGCCCGCGCTTCTCATTTGCCAGCTTTCAGGAGCTGGAGAGCCGCTGGAGCACCC 680
Db 452 GGCCTGCCCGCGCTTCTCATTTGCCAGCTTTCAGGAGCTGGAGAGCCGCTGGAGCACCC 511
QY 681 CTTGACGCGGCGAGCCAGCCACCTGTGATGACCTTCTCATCTCTACTGCGGTGAGCAGAT 740
Db 512 CTTGACGCGGCGAGCCAGCCACCTGTGATGACCTTCTCATCTCTACTGCGGTGAGCAGAT 571
QY 741 CGGACAGAGATCCGCAAGATCACCGGCTGCTTCCACTGCCAGCTTCCCGTTCTGCA 800
Db 572 CGGACAGAGATCCGCAAGATCACCGGCTGCTTCCACTGCCAGCTTCCCGTTCTGCA 631
QY 801 GCAGAGAGAGGCGCGCTCGGGGCTTGCAGAGCTGCAACAGCAGAGCTGCAACAGCAGAGCTGCA 860
Db 632 GCAGAGAGAGGCGCGCTCGGGGCTTGCAGAGCTGCAACAGCAGAGCTGCAACAGCAGAGCTGCA 691
QY 861 GAGAGTCTTCTGGGAGAGACAGAGCGTTCCTGACCAAGTGTCTAGGCGGCTGCTGAGCT 920
Db 692 GAGAGTCTTCTGGGAGAGACAGAGCGTTCCTGACCAAGTGTCTAGGCGGCTGCTGAGCT 751
QY 921 GCTTCCGCGCAGGCGAGTGCAGTCCACAGATGAAGCGGTGTACCTGGCCCTGAACCA 980
Db 752 GCTTCCGCGCAGGCGAGTGCAGTCCACAGATGAAGCGGTGTGTACCTGGCCCTGAACCA 811
QY 981 GTGACGCTGTAGCACACACCAAGTGTCTTATTCAGGAGCTGTGTCTGTGCGAGA 1040
Db 812 GTGACGCTGTAGCACACACCAAGTGTCTTATTCAGGAGCTGTGTCTGTGCGAGA 871
QY 1041 CTTGCCCGCTTGCAGGAGGCGCTTGCAGGACAGCTCGATGAGAGGAGGAGTGAAGTCCCGT 1100
Db 872 CTTGCCCGCTTGCAGGAGGCGCTTGCAGGACAGCTCGATGAGAGGAGGAGTGAAGTCCCGT 931
```

QY 1101 GGCTCAGGCATCCCTGCGGGACATGCCCCACACTCATCGCACCAACCGTTTAC 1160
| | | | |
Db 932 GGTCTACCGCATCCCTGCGGGACATGCCCCACACTCATCGCACCAACCGTTTAC 991
| | | | |
QY 1161 GGCAGAGCTTCAGAGGCATCGTGGATCGCTACGGCGTGGCGCTTACAGAGAGGTCAACCC 1220
| | | | |
Db 992 GGCAGAGCTTCAGAGGCATCGTGGATGCTACGGCGTGGCGCTTACAGAGAGGTCAACCC 1051
| | | | |
QY 1221 CGCTCCCTACACATCATCACTTCCCTTCCCTGCTTGTGCTGTGATGTTTCGGGAGTGGG 1280
| | | | |
Db 1052 CGCTCCCTACACATCATCACTTCCCTTCCCTGCTTGTGCTGTGATGTTTCGGGAGTGGG 1111
| | | | |
QY 1281 CCAGGCGTGTCTCATGTTCCTTTCGCGCTGCGCATGCTTGGGAGAACCGACCGGC 1340
| | | | |
Db 1112 CCAGGCGTGTCTCATGTTCCTTTCGCGCTGCGCATGCTTGGGAGAACCGACCGGC 1171
| | | | |
QY 1341 TGTGAAGCCGCGAGAACAGAGATCTGCGAGACTTCTTACAGGGCGCGCTACCTGCTCCT 1400
| | | | |
Db 1172 TGTGAAGCCGCGAGAACAGAGATCTGCGAGACTTCTTACAGGGCGCGCTACCTGCTCCT 1231
| | | | |
QY 1401 GCTTATGGGCTGTCTTCTCACTACACCGGCTTCACTTACAAACAGAGTCTTCACTGCGC 1460
| | | | |
Db 1232 GCTTATGGGCTGTCTTCTCACTACACCGGCTTCACTTACAAACAGAGTCTTCACTGCGC 1291
| | | | |
QY 1461 CACCAGCATCTTCCCTCGGCTGGAGTGGCGCCATGGCAACAGTCTGCTGGAG 1520
| | | | |
Db 1292 CACCAGCATCTTCCCTCGGCTGGAGTGGCGCCATGGCAACAGTCTGCTGGAG 1351
| | | | |
QY 1521 TGATGCATCTCTGCGCCAGCACAGATGCTTACCCCTGGATCCCAACGTACCGGTGCTT 1580
| | | | |
Db 1352 TGATGCATCTCTGCGCCAGCACAGATGCTTACCCCTGGATCCCAACGTACCGGTGCTT 1411
| | | | |
QY 1581 CTGCGGACCTTACCCCTTTGGCATCGATCTTATTTGGAGCTGTGCTGCCAACCACTTGAG 1640
| | | | |
Db 1412 CTGCGGACCTTACCCCTTTGGCATCGATCTTATTTGGAGCTGTGCTGCCAACCACTTGAG 1471
| | | | |
QY 1641 CTTCCTCAACTCTTCAAGATGAAGATGTCGTGCTGCTGGCGTCTGTCACATGGCCTT 1700
| | | | |
Db 1472 CTTCCTCAACTCTTCAAGATGAAGATGTCGTGCTGCTGGCGTCTGTCACATGGCCTT 1531
| | | | |
QY 1701 TGGGGTGGTCTCTCGAGTCTTTCAACACAGCTGACCTTTGGCCAGAGGACACCGGCTGCT 1760
| | | | |
Db 1532 TGGGGTGGTCTCTCGAGTCTTTCAACACAGCTGACCTTTGGCCAGAGGACACCGGCTGCT 1591
| | | | |
QY 1761 GAGAGCGCTGCGGAGTCACTTCTGCTGGGACTCTTCGGTTTACCTGCTGTTCTTCTAGT 1820
| | | | |
Db 1592 GAGAGCGCTGCGGAGTCACTTCTGCTGGGACTCTTCGGTTTACCTGCTGTTCTTCTAGT 1651
| | | | |
QY 1821 CATCTCAAGTGGCTGTGCTGCGGCTGCGAGGCGGCTCG ---CCCAGCATCTCAT 1877
| | | | |
Db 1652 CATCTCAAGTGGCTGTGCTGCGGCTGCGAGGCGGCTCGGCGCCCGCATCTCAT 1711
| | | | |
QY 1878 CCACCTTCAATCAAGTGTCTTCTTCCACAGCCCGCAGCAACAGGTGCTTACCCCG 1937
| | | | |
Db 1712 CCACCTTCAATCAAGTGTCTTCTTCCACAGCCCGCAGCAACAGGTGCTTACCCCG 1771
| | | | |
QY 1938 GCAGGAGTGGTCCAGGCAACGCTGTGGTCTTGGCTTGGCCATGGTGGCCATCTTCTGCT 1997
| | | | |
Db 1772 GCAGGAGTGGTCCAGGCAACGCTGTGGTCTTGGCTTGGCCATGGTGGCCATCTTCTGCT 1831
| | | | |
QY 1998 GCTTGGCACACCCCTGCACTGTGTCACCGCCACCGCGCGCTGCGGAGGAGGCGCCG 2057
| | | | |
Db 1832 GCTTGGCACACCCCTGCACTGTGTCACCGCCACCGCGCGCTGCGGAGGAGGCGCCG 1891
| | | | |
QY 2058 TGACCGACAGGAGGAGAAACAGGCGCGGTTGCTGGACCTGCTGACGCATCTGTGAATGG 2117
| | | | |
Db 1892 TGACCGACAGGAGGAGAAACAGGCGCGGTTGCTGGACCTGCTGACGCATCTGTGAATGG 1951
| | | | |
QY 2118 CTGGAGCTCCGATCAGGAGAAAGGAGGCGGCTGGATGATGAAGAGGAGGCGCGAGTCTGT 2177
| | | | |
Db 1952 CTGGAGCTCCGATCAGGAGAAAGGAGGCGGCTGGATGATGAAGAGGAGGCGCGAGTCTGT 2011
| | | | |
QY 2178 CCCTCGAGGTGCTCATGTGACACAGGCCATCCACACCATCGAGTTCTGCTGGGCTGCGT 2237
| | | | |

Db 2012 CCCCTCCGAGGTGCTCATGTGACACGAGGCATCCACACCATCGAGTTCTGCTGGGCTGCGT 2071
| | | | |
QY 2238 CTCCAAACAACCGCTTCTTACCTGCGCTTGTGGGCTTGTAGCCTTGGCCACGCCCAGCTGTC 2297
| | | | |
Db 2072 CTCCAAACAACCGCTTCTTACCTGCGCTTGTGGGCTTGTAGCCTTGGCCACGCCCAGCTGTC 2131
| | | | |
QY 2298 CGAGGTTCTGTGGGCTTGTGATGCGCATAGGCTTGGGCTTGGGCGGAGGTGGGCT 2357
| | | | |
Db 2132 CGAGGTTCTGTGGGCTTGTGATGCGCATAGGCTTGGGCTTGGGCGGAGGTGGGCT 2191
| | | | |
QY 2358 GCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2417
| | | | |
Db 2192 GCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2251
| | | | |
QY 2418 GCTGCTGATGAGGAGGACTCTCAGCCTTCTGCGCGCTTGGGCTGCGCTGCGCTGCGCTGCGCT 2477
| | | | |
Db 2252 GCTGCTGATGAGGAGGACTCTCAGCCTTCTGCGCGCTTGGGCTGCGCTGCGCTGCGCTGCGCT 2311
| | | | |
QY 2478 CCAGAACAAAGTTCTTACTCAGGACGCGCTACAGCTGAGTCCCTTCACTTCTGCTGCGCT 2537
| | | | |
Db 2312 CCAGAACAAAGTTCTTACTCAGGACGCGCTACAGCTGAGTCCCTTCACTTCTGCTGCGCT 2371
| | | | |
QY 2538 AGATGACTAGGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCT 2597
| | | | |
Db 2372 AGATGACTAGGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCT 2431
| | | | |
QY 2598 AGGAATAAAGACGCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCT 2640
| | | | |
Db 2432 AGGAATAAAGACGCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCTTCTGCGGCT 2474
| | | | |

Search completed: June 30, 2006, 10:38:33

Job time : 2798 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2006, 06:17:20 ; Search time 307 Seconds
(without alignments)
10168.695 Million cell updates/sec

Title: US-10-783-519-1

Perfect score: 2640

Sequence: 1 cggcgcgcgagcgagcag.....gcaaaaaaaaaaaaaa 2640

Scoring table:

OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 809770 seqs, 591248006 residues

Word size : 1

Total number of hits satisfying chosen parameters: 1619416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA New:

- 1: /EMC Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB.seq:
- 2: /EMC Celerra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB.seq:
- 3: /EMC Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB.seq:
- 4: /EMC Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB.seq:
- 5: /EMC Celerra_SIDS3/ptodata/2/pubpna/PCT_NEW_PUB.seq:
- 6: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB.seq:
- 7: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq:
- 8: /EMC Celerra_SIDS3/ptodata/2/pubpna/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-----------------------|
| 1 | 2142 | 81.1 | 2700 | 6 | US-10-511-937-410 |
| 2 | 775 | 29.4 | 1696 | 7 | US-11-266-748A-76629 |
| 3 | 775 | 29.4 | 1696 | 7 | US-11-266-748A-109489 |
| 4 | 775 | 29.4 | 1696 | 7 | US-11-266-748A-129440 |
| 5 | 760 | 28.8 | 1000 | 7 | US-11-266-748A-221798 |
| 6 | 760 | 28.8 | 1000 | 7 | US-11-266-748A-287153 |
| 7 | 760 | 28.8 | 1000 | 7 | US-11-266-748A-338582 |
| 8 | 760 | 28.8 | 1000 | 7 | US-11-266-748A-397768 |
| 9 | 760 | 28.8 | 1000 | 7 | US-11-266-748A-468814 |
| 10 | 581 | 22.0 | 1042 | 7 | US-11-266-748A-76627 |
| 11 | 581 | 22.0 | 1042 | 7 | US-11-266-748A-109487 |
| 12 | 581 | 22.0 | 1042 | 7 | US-11-266-748A-129438 |
| 13 | 532 | 20.2 | 671 | 7 | US-11-266-748A-54130 |
| 14 | 500 | 18.9 | 738 | 7 | US-11-266-748A-76628 |
| 15 | 500 | 18.9 | 738 | 7 | US-11-266-748A-109488 |
| 16 | 500 | 18.9 | 738 | 7 | US-11-266-748A-129439 |
| 17 | 489 | 18.5 | 831 | 7 | US-11-266-748A-8067 |
| 18 | 454 | 17.2 | 561 | 7 | US-11-266-748A-362824 |
| 19 | 454 | 17.2 | 561 | 7 | US-11-266-748A-446203 |
| 20 | 423 | 16.0 | 1839 | 7 | US-11-266-748A-361154 |
| 21 | 423 | 16.0 | 1839 | 7 | US-11-266-748A-444533 |
| 22 | 391 | 14.8 | 1469 | 7 | US-11-266-748A-98830 |
| 23 | 391 | 14.8 | 1469 | 7 | US-11-266-748A-151641 |
| 24 | 340 | 12.9 | 1127 | 7 | US-11-266-748A-76630 |
| 25 | 340 | 12.9 | 1127 | 7 | US-11-266-748A-109490 |

| | | | | | | | |
|---|----|-----|------|------|---|-----------------------|-------------------|
| c | 26 | 340 | 12.9 | 1127 | 7 | US-11-266-748A-129441 | Sequence 129441, |
| c | 27 | 227 | 8.6 | 705 | 7 | US-11-266-748A-367398 | Sequence 367398, |
| c | 28 | 227 | 8.6 | 705 | 7 | US-11-266-748A-450777 | Sequence 450777, |
| c | 29 | 196 | 7.4 | 884 | 7 | US-11-266-748A-76631 | Sequence 76631, A |
| c | 30 | 196 | 7.4 | 884 | 7 | US-11-266-748A-109491 | Sequence 109491, |
| c | 31 | 196 | 7.4 | 884 | 7 | US-11-266-748A-129442 | Sequence 129442, |
| c | 32 | 50 | 1.9 | 50 | 6 | US-10-511-937-78 | Sequence 78, Appl |
| c | 33 | 33 | 1.2 | 446 | 6 | US-10-488-619-1100 | Sequence 1100, Ap |
| c | 34 | 32 | 1.2 | 3137 | 7 | US-11-266-748A-30495 | Sequence 30495, A |
| c | 35 | 26 | 1.0 | 571 | 6 | US-10-488-619-2078 | Sequence 2078, Ap |
| c | 36 | 26 | 1.0 | 597 | 6 | US-10-488-619-1857 | Sequence 1857, Ap |
| c | 37 | 26 | 1.0 | 4692 | 6 | US-10-511-937-646 | Sequence 646, App |
| c | 38 | 26 | 1.0 | 4782 | 7 | US-11-266-748A-31602 | Sequence 31602, A |
| c | 39 | 25 | 0.9 | 444 | 6 | US-10-488-619-1101 | Sequence 1101, Ap |
| c | 40 | 25 | 0.9 | 1000 | 7 | US-11-266-748A-115175 | Sequence 115175, |
| c | 41 | 25 | 0.9 | 1000 | 7 | US-11-266-748A-157339 | Sequence 157339, |
| c | 42 | 25 | 0.9 | 1000 | 7 | US-11-266-748A-283882 | Sequence 283882, |
| c | 43 | 25 | 0.9 | 1000 | 7 | US-11-266-748A-335311 | Sequence 335311, |
| c | 44 | 25 | 0.9 | 1000 | 7 | US-11-266-748A-393689 | Sequence 393689, |
| c | 45 | 25 | 0.9 | 1000 | 7 | US-11-266-748A-464735 | Sequence 464735, |

ALIGNMENTS

RESULT 1

US-10-511-937-410
; Sequence 410, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 410
; LENGTH: 2700
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-511-937-410

| | | | | |
|-----------------------|-----------------|---|-----------|--------------|
| Query Match | 81.1%; | Score 2142; | DB 6; | Length 2700; |
| Best Local Similarity | 99.7%; | Pred. No. 0; | | |
| Matches 2582; | Conservative 0; | Mismatches 4; | Indels 3; | Gaps 1; |
| QY | 38 | GCACGACACCCCGGGACCATGGCTCCATGTTCCGAGCAGAGAGTGGCCCTGTGCC | 97 | |
| Db | 89 | GCACGACACACCCCGGGACCATGGCTCCATGTTCCGAGCAGAGAGTGGCCCTGTGCC | 148 | |
| QY | 98 | AGCTCTTCTGCGCCACAGCGGCTTACACCTCGGTGAGTCCGCTCGGCGAGCTGGGCC | 157 | |
| Db | 149 | AGCTCTTCTGCGCCACAGCGGCTTACACCTCGGTGAGTCCGCTCGGCGAGCTGGGCC | 208 | |
| QY | 158 | TCGTGGAGTTTCAGAGACCTCAACGCTTCGCTGAGCGCTTCCAGAGACGCTTTGTGTTG | 217 | |
| Db | 209 | TCGTGGAGTTTCAGAGACCTCAACGCTTCGCTGAGCGCTTCCAGAGACGCTTTGTGTTG | 268 | |

| | | | |
|----|------|---|------|
| QY | 218 | ATGTTTGGCGCTGTGAGAGCTGGAGAAAGACTTCACTTCTGCAGGAGGAGTGC | 277 |
| DB | 269 | ATGTTTCGGCGCTGTGAGAGCTGGAGAAAGACTTCACTTCTGCAGGAGGAGTGC | 328 |
| QY | 278 | GGGCTTGGGCTGGTCTGTGCCCCCGCCAAAGGGGAGGCTGCCGGCACCCCCACCCCGGGAC | 337 |
| DB | 329 | GGGCTTGGGCTGGTCTGTGCCCCCGCCAAAGGGGAGGCTGCCGGCACCCCCACCCCGGGAC | 388 |
| QY | 338 | TGCTGCGCATTCAGAGAGAGACGGAGCGCTTGGCCCCAGAGAGCTGCGGGATGTGGGGGACA | 397 |
| DB | 389 | TGCTGCGCATTCAGAGAGAGAGAGAGCGCTTGGCCCCAGAGAGCTGCGGGATGTGGGGGACA | 448 |
| QY | 398 | ACCAGAGCGCTTCGGGCCCCAGCTGCACCAAGCTGCAGCTCCACGCCCGCTGTACGCC | 457 |
| DB | 449 | ACCAGAGCGCTTCGGGCCCCAGCTGCACCAAGCTGCAGCTCCACGCCCGCTGTACGCC | 508 |
| QY | 458 | AGGGCCATGAACCTTCAGCTGGCAGCGGCCACACAGATGGGGCTTCAGAGAGAGCGCCC | 517 |
| DB | 509 | AGGGCCATGAACCTTCAGCTGGCAGCGGCCACACAGATGGGGCTTCAGAGAGAGCGCCC | 568 |
| QY | 518 | TGCTCCAGGCCCCCGGGGGCGCACACAGAGACCTGAGGCTCAACTTTGTGTGACGGTCCG | 577 |
| DB | 569 | TGCTCCAGGCCCCCGGGGGCGCACACAGAGACCTGAGGCTCAACTTTGTGTGACGGTCCG | 628 |
| QY | 578 | TGGAGCCCCACAAGGCCCTTCCTCTAGAGCGCTGCTTGTGAGGGCTTGC | 637 |
| DB | 629 | TGGAGCCCCACAAGGCCCTTCCTCTAGAGCGCTGCTTGTGAGGGCTTGC | 688 |
| QY | 638 | TCATTGCCAGCTTCAGGGAGCTGGAGAGCGCTTGGAGACACCCGCTGACCGGGCAGCAG | 697 |
| DB | 689 | TCATTGCCAGCTTCAGGGAGCTGGAGAGCGCTTGGAGACACCCGCTGACCGGGCAGCAG | 748 |
| QY | 698 | CCAGTGGATGACCTTCCTCATCTCTACTTGGGGTGAGCAGATCGGACGAGAGATCCGCA | 757 |
| DB | 749 | CCAGTGGATGACCTTCCTCATCTCTACTTGGGGTGAGCAGATCGGACGAGAGATCCGCA | 808 |
| QY | 758 | AGATCACGGACTGCTTCCACTGCCCCACGCTTTCGCCGTTTCTGACGACGAGAGCGCCGCC | 817 |
| DB | 809 | AGATCACGGACTGCTTCCACTGCCCCACGCTTTCGCCGTTTCTGACGACGAGAGCGCCGCC | 868 |
| QY | 818 | TCGGGGCCCTGCAGAGCTGCAACAGACAGACAGCAGCTGCAGGAGGTTCTTCGGGGAGA | 877 |
| DB | 869 | TCGGGGCCCTGCAGAGCTGCAACAGACAGACAGCAGCTGCAGGAGGTTCTTCGGGGAGA | 928 |
| QY | 878 | CAGAGCGGTTCTTGAGCCAGGTGCTAGGCCGGGTGCTGACGCTGCTGCCACAGGGCAGG | 937 |
| DB | 929 | CAGAGCGGTTCTTGAGCCAGGTGCTAGGCCGGGTGCTGACGCTGCTGCCACAGGGCAGG | 988 |
| QY | 938 | TGCAGGTCCACAAGATGAAGCCCGTGTACTTGGCCCTGAAACAGTCAGCGGTGAGCACCA | 997 |
| DB | 989 | TGCAGGTCCACAAGATGAAGCCCGTGTACTTGGCCCTGAAACAGTCAGCGGTGAGCACCA | 1048 |
| QY | 998 | CGCACAGTGCCTCATTTGCCAGAGCGCTGTGCTCTGTGCGAGACCTGCCCGCCTTGAGG | 1057 |
| DB | 1049 | CGCACAGTGCCTCATTTGCCAGAGCGCTGTGCTCTGTGCGAGACCTGCCCGCCTTGAGG | 1108 |
| QY | 1058 | AGGCCCTTGCGGGACAGCTTCGATGAGAGAGGAGTGAGTGGCGTGCTCACCGATCCCTCT | 1117 |
| DB | 1109 | AGGCCCTTGCGGGACAGCTTCGATGAGAGAGGAGTGAGTGGCGTGCTCACCGATCCCTCT | 1168 |
| QY | 1118 | GCCGGGACATGCCCCCACA | 1177 |
| DB | 1169 | GCCGGGACATGCCCCCACA | 1228 |
| QY | 1178 | TCGTGGATTCGCTACGGCGTGCGGCGCTTACAGAGAGGTCAACCCCGCTTCCCTTACACATCA | 1237 |
| DB | 1229 | TCGTGGATTCGCTACGGCGTGCGGCGCTTACAGAGAGGTCAACCCCGCTTCCCTTACACATCA | 1288 |
| QY | 1238 | TCACCTTCCCCTTCTGTTTCTGTGATGTTTGGGGATGTGGGCCACGGGGCTGCTCATGT | 1297 |
| DB | 1289 | TCACCTTCCCCTTCTGTTTCTGTGATGTTTGGGGATGTGGGCCACGGGGCTGCTCATGT | 1348 |
| QY | 1298 | TCCTCTTTGCCCTCGCCATGGTCTCTTCGGAGAAACCGACCGGCTGTGAAAGCGCGCAGA | 1357 |

| | | | |
|----|------|---|------|
| Db | 1349 | TCCTCTTCGGCCCTGGCCATGGTCTCTTGGGAGAAACGACCGGCTGTGAAGGCGCGCAGA | 1408 |
| Qy | 1358 | ACGAGATCTGGCAGACTTCTTTCAGGGGCGGTACCTGTCTCTGCTTATGGGCGCTGTCT | 1417 |
| Db | 1409 | ACGAGATCTGGCAGACTTCTTTCAGGGGCGGTACCTGTCTCTGCTTATGGGCGCTGTCT | 1468 |
| Qy | 1418 | CCATCTACACGGCTTCATCTACAACAGATGCTTCAGTTCGGCGCACACAGCATCTTCCCT | 1477 |
| Db | 1469 | CCATCTACACGGCTTCATCTACAACAGATGCTTCAGTTCGGCGCACACAGCATCTTCCCT | 1528 |
| Qy | 1478 | CGGCTGGAGTGTGGCGCCGATGGCCAAACAGTCTGGCTGGAGTGATGCAATTCCTGGCCC | 1537 |
| Db | 1529 | CGGCTGGAGTGTGGCGCCGATGGCCAAACAGTCTGGCTGGAGTGATGCAATTCCTGGCCC | 1588 |
| Qy | 1538 | AGCACACGATGCTTACCTTGGATGCCAAACAGTCTGGCTGGAGTGATGCAATTCCTGGCCC | 1597 |
| Db | 1589 | AGCACACGATGCTTACCTTGGATGCCAAACAGTCTGGCTGGAGTGATGCAATTCCTGGCCC | 1648 |
| Qy | 1598 | TTGGCATCGATCTTATTTGGAGCTGGCTGCCAAACACTTGGCTTCTCAACTCTCTTCA | 1657 |
| Db | 1649 | TTGGCATCGATCTTATTTGGAGCTGGCTGCCAAACACTTGGCTTCTCAACTCTCTTCA | 1708 |
| Qy | 1658 | AGATGAAGATGTCCGTCACTCTGGGCGTGGTGACATGGCTTTGGGGTGGTCTCGGAG | 1717 |
| Db | 1709 | AGATGAAGATGTCCGTCACTCTGGGCGTGGTGACATGGCTTTGGGGTGGTCTCGGAG | 1768 |
| Qy | 1718 | TCTTCAACACGTGCACTTTGGCCAGAGGCAACGGCTGTCTGGAGAGACGTGCCGGAGC | 1777 |
| Db | 1769 | TCTTCAACACGTGCACTTTGGCCAGAGGCAACGGCTGTCTGGAGAGACGTGCCGGAGC | 1828 |
| Qy | 1778 | TCACCTTCCTGCTGGGACTCTTCGGTTACTCTGGTTCTAGTCACTCAAGTGCTGT | 1837 |
| Db | 1829 | TCACCTTCCTGCTGGGACTCTTCGGTTACTCTGGTTCTAGTCACTCAAGTGCTGT | 1888 |
| Qy | 1838 | GTGTCTGGGCTGCCAGGGCGCGCTCG --- CCCAGCATCCTCATPCCAATTCATCAACATGT | 1894 |
| Db | 1889 | GTGTCTGGGCTGCCAGGGCGCGCTCGGCCCCAGCATCCTCATCCAATTCATCAACATGT | 1948 |
| Qy | 1895 | TCCTCTTCTCCACAGGCCAGCAACAGGTGTCTTACCCCGGCAAGAGTGGTCCAGG | 1954 |
| Db | 1949 | TCCTCTTCTCCACAGGCCAGCAACAGGTGTCTTACCCCGGCAAGAGTGGTCCAGG | 2008 |
| Qy | 1955 | CCAGCTGTGTGCTCTGGCCCTGGCCATGTGGCCCATGTGGTCCCATGTGGTGGCAACCCCTGC | 2014 |
| Db | 2009 | CCAGCTGTGTGCTCTGGCCCTGGCCATGTGGCCCATGTGGTGGTGGCAACCCCTGC | 2068 |
| Qy | 2015 | ACCTGTGTGCAACGCCACCGCGCGCTGGAGGAGGCCGCTGACCGACAGGAGGAAA | 2074 |
| Db | 2069 | ACCTGTGTGCAACGCCACCGCGCGCTGGAGGAGGCCGCTGACCGACAGGAGGAAA | 2128 |
| Qy | 2075 | ACAAGGCCGGTGTGCGACCTGCTGACGCACTGTGTAATGGCTGGAGCTCCGATGAGG | 2134 |
| Db | 2129 | ACAAGGCCGGTGTGCGACCTGCTGACGCACTGTGTAATGGCTGGAGCTCCGATGAGG | 2188 |
| Qy | 2135 | AAAAGGCAGGGGCTTGGATGATGAAGAGAGAGCGGAGCTCGTCCCTCCGAGGTCTCTCA | 2194 |
| Db | 2189 | AAAAGGCAGGGGCTTGGATGATGAAGAGAGAGCGGAGCTCGTCCCTCCGAGGTCTCTCA | 2248 |
| Qy | 2195 | TGCAACAGGCCATCCACACCATCAGTTCTGCTGGGCTGGCTCTCCAAACACCGCTCTCT | 2254 |
| Db | 2249 | TGCAACAGGCCATCCACACCATCAGTTCTGCTGGGCTGGCTCTCCAAACACCGCTCTCT | 2308 |
| Qy | 2255 | ACCTGGCTGTGGGCCCTGAGCCTGGCCGACCGCCAGCTGTCCGAGGTTTGTGTGGGCCA | 2314 |
| Db | 2309 | ACCTGGCTGTGGGCCCTGAGCCTGGCCGACCGCCAGCTGTCCGAGGTTTGTGTGGGCCA | 2368 |
| Qy | 2315 | TGTGTATGCCATAGCCTGGGCTGGCGGGAGGTGGCGTGGCGGCTGTGGTCTCTGG | 2374 |
| Db | 2369 | TGTGTATGCCATAGCCTGGGCTGGCGGGAGGTGGCGTGGCGGCTGTGGTCTCTGG | 2428 |
| Qy | 2375 | TCGCCATCTTTGGCGCTTTGGCGTGATACCGTGGCTATCTCTGTCTGTGTATGGAGGGAC | 2434 |

| | | | |
|----|------|---|------|
| Db | 2429 | TCCCATCTTTGGCGCCTTTGCCCGTGATGACCGGTGGCTATCCTCTCGTGGTATGAGGGAC | 2488 |
| Qy | 2435 | TCTCAGCCTTCTCTGACGCCCTGGCGGTGCACTGGGTGGAAATCCAGAAACAAGTTCTTACT | 2494 |
| Db | 2489 | TCTCAGCCTTCTCTGACGCCCTGGCGGTGCACTGGGTGGAAATCCAGAAACAAGTTCTTACT | 2548 |
| Qy | 2495 | CAGGCACGGGGCTACAAGCTGAGTCCCTTTCACCTTCGCTGCCACAGATGACTAGGGCCCCAC | 2554 |
| Db | 2549 | CAGGCACGGGGCTACAAGCTGAGTCCCTTTCACCTTCGCTGCCACAGATGACTAGGGCCCCAC | 2608 |
| Qy | 2555 | TGCAGGTCTCTGCCACAGACTCTCTTCTGACCTCTCAGGCAGGAGAGGAAATAAAGACGGTCC | 2614 |
| Db | 2609 | TGCAGGTCTCTGCCACAGACTCTCTTCTGACCTCTGAGGCAGGAGAGGAAATAAAGACGGTCC | 2668 |
| Qy | 2615 | GCCTTGGCA | 2623 |
| Db | 2669 | GCCTTGGCA | 2677 |

RESULT 2

US-11-266-748A-76629
; Sequence 76629, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76629
; LENGTH: 1696
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-76629

| Query Match | 29.4% | Score 775; | DB 7; | Length 1696; | |
|-----------------------|-----------------|---------------------|-----------|---|------|
| Best Local Similarity | 99.3%; | Pred. No. 1.2e-212; | | | |
| Matches 1505; | Conservative 0; | Mismatches 5; | Indels 5; | Gaps 3; | |
| Qy 781 | CACGTC | TTCCG | TTTCTG | CAGCAGGAGGAGCCGCTCGGGGCCCTGACGAGCTGCA | 840 |
| | | | | | |
| Db 28 | CACGTC | TTCCG | TTTCTG | CAGCAGGAGGAGCCGCTCGGGGCCCTGACGAGCTGCA | 87 |
| | | | | | |
| Qy 841 | CAGCAG | CCAGG | AGCTG | CAGGAGGTCTCGGGGAGACAGAGCGGTTCTTGAGCAGG | 900 |
| | | | | | |
| Db 88 | CAGCAG | CCAGG | AGCTG | CAGGAGGTCTCGGGGAGACAGAGCGGTTCTTGAGCAGG | 147 |
| | | | | | |
| Qy 901 | CTAGG | CCGGT | GTGTC | GAGCTGCTGCCCGCAGGCGAGGTGCAGGTCCACAGATGAAG | 960 |
| | | | | | |
| Db 148 | CTAGG | CCGGT | GTGTC | GAGCTGCTGCCCGCAGGCGAGGTGCAGGTCCACAGATGAAG | 207 |
| | | | | | |
| Qy 961 | GTGTAC | CTGCC | CTGTGA | ACCACTGTCAGCGGTGAGCACCAAGTGCCTCATTTGCCGAG | 1020 |
| | | | | | |

[illegible]

Db 1288 GCCGCTCGGAGGAGCCGCTGACCGACAGGAGAAACAAAGCCGGCTTGCTGGACC 1347
Qy 2096 TGCCTGACGCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGGAGGGGCTGGATG 2155
Db 1348 TGCCTGACGCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGGAGGGGCTTGATG 1407
Qy 2156 ATGAAGAGGAGGCGGAGCTGCTCCCTCCGAGGTGCTCATGACACAGGCGCATCCACACCA 2215
Db 1408 ATGAAGAGGAGGCGGAGCTGCTCCCTCCGAGGTGCTCATGACACAGGCGCATCCACACCA 1467
Qy 2216 TCGAGTTCCTGCTGGGCTGCTTCCAAACACGCGCTTCTTACCTGCGCCTGTGGGCCCTGA 2275
Db 1468 TCGAGTTCCTGCTGGGCTGCTTCCAAACACGCGCTTCTTACCTGCGCCTGTGGGCCCTGA 1527
Qy 2276 GCGTGGCCCAAGGCC 2290
Db 1528 GCGTGGCCCAAGGCC 1542

RESULT 3

US-11-266-748A-109489
; Sequence 109489, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnneton, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 109489
; LENGTH: 1696
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-109489

Query Match 29.4%; Score 775; DB 7; Length 1696;
Best Local Similarity 99.3%; Pred. No. 1.2e-212;
Matches 1505; Conservative 0; Mismatches 5; Indels 5; Gaps 3;

Qy 781 CACGTCTCCCGTTTCTGACGAGGAGGAGGCGCGCTCGGGGCGCTGACGAGCTGCA 840
Db 28 CACGTCTCCCGTTTCTGACGAGGAGGAGGCGCGCTCGGGGCGCTGACGAGCTGCA 87
Qy 841 CAGCAGGACGAGGCTGACGAGGCTTCTCGGGAGACAGAGCGGTTCCTGAGCCAGGTG 900
Db 88 CAGCAGGACGAGGCTGACGAGGCTTCTCGGGAGACAGAGCGGTTCCTGAGCCAGGTG 147
Qy 901 CTAGGCGGGGTGCTGAGCTGCTGCCGACGAGGAGGCTCCACAGATGAAGGCC 960
Db 148 CTAGGCGGGGTGCTGAGCTGCTGCCGACGAGGAGGCTCCACAGATGAAGGCC 207
Qy 961 GTGTACCTGGCCCTGAACCAAGTGCAGCTGACGACCAACGAGTGCCTATTGCCGAG 1020

Db 208 GTGTACTGGCCCTGAAACGAGTGCAGCTGACGACCAACGACAGTGCCTCATTTGCCGAG 267
Qy 1021 GCCTGGTGTCTGTGCGAGACCTGCCGCCCTGCGAGGAGGCTTGCAGGACAGCTCGATG 1080
Db 268 GCCTGGTGTCTGTGCGAGACCTGCCGCCCTGCGAGGAGGCTTGCAGGACAGCTCGATG 327
Qy 1081 GAGGAGGAGTGAAGTGGCTGCTCACGCACTCCCTGCGGGACATGCCCCCACACTC 1140
Db 328 GAGGAGGAGTGAAGTGGCTGCTCACGCACTCCCTGCGGGACATGCCCCCACACTC 387
Qy 1141 ATCCGACCAACCGCTTACGCGCCAGCTTTCAGGGGATCGTGGATCGCTACGCGGTGGG 1200
Db 388 ATCCGACCAACCGCTTACGCGCCAGCTTTCAGGGGATCGTGGATCGCTACGCGGTGGG 447
Qy 1201 CGCTACGAGGAGTCAACCCCGCTCCCTACACCATCATCCTTCCCTTCTCTGTTTGT 1260
Db 448 CGCTACGAGGAGTCAACCCCGCTCCCTACACCATCATCCTTCCCTTCTCTGTTTGT 507
Qy 1261 GTGATGTTGGGGATGTGGGCCACGGGCTGCTCATGTTCTTTCGCCCTGGCCATGTC 1320
Db 508 GTGATGTTGGGGATGTGGGCCACGGGCTGCTCATGTTCTTTCGCCCTGGCCATGTC 567
Qy 1321 CTTGGGAGAACCGGCTGTGAAAGCCGCGCAGAACGAGATCTGGCAGACTTTTCTTC 1380
Db 568 CTTGGGAGAACCGGCTGTGAAAGCCGCGCAGAACGAGATCTGGCAGACTTTTCTTC 627
Qy 1381 AGGGGCGCTACTGCTCTCTTATGAGGCTGTTCCTCATCTACACCGGCTTATCTATC 1440
Db 628 AGGGGCGCTACTGCTCTCTTATGAGGCTGTTCCTCATCTACACCGGCTTATCTATC 687
Qy 1441 AACGAGTGTCTCAGTGGGCGCACAGCATCTTCCCTCGGGCTGGAGTGTGCGGCCATG 1500
Db 688 AACGAGTGTCTCAGTGGGCGCACAGCATCTTCCCTCGGGCTGGAGTGTGCGGCCATG 747
Qy 1501 GCCAACCAAGTCTGGGCTGAGTGAATTCCTGGGCGCAGCACAGATGTTTACCCCTGGAT 1560
Db 748 GCCAACCAAGTCTGGGCTGAGTGAATTCCTGGGCGCAGCACAGATGTTTACCCCTGGAT 807
Qy 1561 CCCAACGTCACCGGTGTCTTCTCGGAGCCCTAACCCCTTTGGCATCGATCTTATTTGAGC 1620
Db 808 CCCAACGTCACCGGTGTCTTCTCGGAGCCCTAACCCCTTTGGCATCGATCTTATTTGAGC 867
Qy 1621 CTGGGTCGCCAACCACTTGAAGTTCCTCAACTCTTCAAGATGAAGTGTCCGTATCTCTG 927
Db 868 CTGGGTCGCCAACCACTTGAAGTTCCTCAACTCTTCAAGATGAAGTGTCCGTATCTCTG 972
Qy 1681 GCGGTCTGTCACATGGCC-TTTTGGGGT-GGTCTCGGAGTCTTCAACCAAGTGTGCTT 1738
Db 928 GCGGTCTGTCACATGGCC-TTTTGGGGTGGGTCTCTCGGAGTCTTCAACCAAGTGTGCTT 987
Qy 1739 GCCAGAGGACCGGCTGCTGCTGAGAGCGCTGCGGAGCTCACCTTCTGCTGGGACTCT 1798
Db 988 GCCAGAGGACCGGCTGCTGCTGAGAGCGCTGCGGAGCTCACCTTCTGCTGGGACTCT 1047
Qy 1799 TCGGTTACTGCTGTTCTAGTCACTACAGTGGCTGTGTCTGCGGCTGCGGCTGCGAGGCC 1858
Db 1048 TCGGTTACTGCTGTTCTAGTCACTACAGTGGCTGTGTCTGCGGCTGCGGCTGCGAGGCC 1107
Qy 1859 CCTCG---CCCAGCATCTCTCATCCACTTCAATCAAGTGTCTTCTTCTCCACAGCCCCA 1915
Db 1108 CTTGGGCCCCCAGCATCTCTCATCCACTTCAATCAAGTGTCTTCTTCTCCACAGCCCCA 1167
Qy 1916 GCACAGGCTGTCTTACCCCGGAGGAGTGTGTCAGGCGACGCTGTGTGTCTGGCCT 1975
Db 1168 GCACAGGCTGTCTTACCCCGGAGGAGTGTGTCAGGCGACGCTGTGTGTCTGGCCT 1227
Qy 1976 TGGCCATGTGTCCTCTGCTGCTTGGCACACCTCTGACCTGTGACCGCCACCGCC 2035
Db 1228 TGGCCATGTGTCCTCTGCTGCTTGGCACACCTCTGACCTGTGACCGCCACCGCC 1287
Qy 2036 GCGCCTGCGGAGGAGGCGCTGACCGACAGGAGAAACAAAGGCGCGGTTGCTGGACC 2095

Db 1288 GCCGCTGCGAGGAGGCGCTGACCGACAGGAGGAAAAAAGGCGGGTTGCTGGACC 1347
Qy 2096 TGCCTGAGCGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAAAGGCGGGGCTTGGATG 2155
Db 1348 TGCCTGAGCGATCTGTGAATGGCTGGAGCTCCGATGAGGAAAAAGGCGGGGCTTGGATG 1407
Qy 2156 ATGAAGAGGAGGCGGAGCTGCTCCCTCCGAGGTGCTCATCACAGGCGCATCCACACCA 2215
Db 1408 ATGAAGAGGAGGCGGAGCTGCTCCCTCCGAGGTGCTCATCACAGGCGCATCCACACCA 1467
Qy 2216 TCGAGTTTCTGCTGGGCTGCTTCCCAACACCGCTCTTACCTGCGCTGTGGGCGCTGA 2275
Db 1468 TCGAGTTTCTGCTGGGCTGCTTCCCAACACCGCTCTTACCTGCGCTGTGGGCGCTGA 1527
Qy 2276 GCCTGGGCCAGGCC 2290
Db 1528 GCCTGGGCCAGGCC 1542

RESULT 4

US-11-266-748A-129440/c
; Sequence 129440, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 129440
; LENGTH: 1696
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-129440

Query Match 29.4%; Score 775; DB 7; Length 1696;

Best Local Similarity 99.3%; Pred. No. 1.2e-212;

Matches 1505; Conservative 0; Mismatches 5; Indels 5; Gaps 3;

Qy 781 CAGCTCTCCGTTTCTGACGAGGAGGCGCGCTCGGGGCTTGCAGCAGCTCAA 840
Db 1669 CAGCTCTCCGTTTCTGACGAGGAGGCGCGCTCGGGGCTTGCAGCAGCTCAA 1610
Qy 841 CAGCAGAGCCAGGAGCTGCGAGGAGTCTTCGGGAGACAGAGCGTTTCTGAGCCAGTG 900
Db 1609 CAGCAGAGCCAGGAGCTGCGAGGAGTCTTCGGGAGACAGAGCGTTTCTGAGCCAGTG 1550
Qy 901 CTAGGCGGGTGTGAGCTGTGCGCCAGGCGAGGTGCAAGATGAAGGCC 960
Db 1549 CTAGGCGGGTGTGAGCTGTGCGCCAGGCGAGGTGCAAGATGAAGGCC 1490
Qy 961 GTGTACTGGCCCTGAACACAGTGCAGCGTGAGCAGCAGGCAAGTGCCTATTGGCCAG 1020

Db 1489 GTGTACCTGGCCCTGAACCAAGTGCAGCGTGAAGCAAGCGCAAGATGCTCTATTGCCGAG 1430
Qy 1021 GCCTGGTGTCTGTGCGAGAGCTTGCCTCCGCTGAGAGGCGCTGCGGAGACAGCTCGATG 1080
Db 1429 GCCTGGTGTCTGTGCGAGAGCTTGCCTCCGCTGAGAGGCGCTGCGGAGACAGCTCGATG 1370
Qy 1081 GAGAGGAGTGTGAGTGCCTGGCTCAACCGCATTCCTTGCCTGCGGAGACATGCCCCCACTC 1140
Db 1369 GAGAGGAGTGTGAGTGCCTGGCTCAACCGCATTCCTTGCCTGCGGAGACATGCCCCCACTC 1310
Qy 1141 ATCCGACCAACCGCTTACGCGGCGCTTCCAGGCGCATTCGAGGCGATCGTTCGCTGGCG 1200
Db 1309 ATCCGACCAACCGCTTACGCGGCGCTTCCAGGCGCATTCGAGGCGATCGTTCGCTGGCG 1250
Qy 1201 CGCTACGAGGAGTCAACCCCGCTTCCATACCATCATCATCATCATCATCATCATCATCAT 1260
Db 1249 CGCTACGAGGAGTCAACCCCGCTTCCATACCATCATCATCATCATCATCATCATCATCAT 1190
Qy 1261 GTGATGTTGGGGATGTGGGCGACGCGGTGCTCATGTTCCTTCTGCGCTTGGCCATGCTC 1320
Db 1189 GTGATGTTGGGGATGTGGGCGACGCGGTGCTCATGTTCCTTCTGCGCTTGGCCATGCTC 1130
Qy 1321 CTTGCGGAGAACCGACCGGCTGTGAAGCGCGCGAGAACGAGATCTGCGACACTTCTTC 1380
Db 1129 CTTGCGGAGAACCGACCGGCTGTGAAGCGCGCGAGAACGAGATCTGCGACACTTCTTC 1070
Qy 1381 AGGGCGCGCTACCTGCTTCTGCTTATGGGCTGTTCCTCATCTACACCGGCTTCACTAC 1440
Db 1069 AGGGCGCGCTACCTGCTTCTGCTTATGGGCTGTTCCTCATCTACACCGGCTTCACTAC 1010
Qy 1441 AACGAGTGTTCAGTTCGCGGCGACAGCATCTTCCCTCGGCTGGAGTGTGGCGCCATG 1500
Db 1009 AACGAGTGTTCAGTTCGCGGCGACAGCATCTTCCCTCGGCTGGAGTGTGGCGCCATG 950
Qy 1501 GCCAACAGTCTGGCTGGAGTGTGATTCCTGCGCGGCGAGACAGATCTTACCTGAGT 1560
Db 949 GCCAACAGTCTGGCTGGAGTGTGATTCCTGCGCGGCGAGACAGATCTTACCTGAGT 890
Qy 1561 CCCAACGTCACCGGTGTCTTCTGGGACCTTACCCCTTGGCATCGATCTTATTTGGAGC 1620
Db 889 CCCAACGTCACCGGTGTCTTCTGGGACCTTACCCCTTGGCATCGATCTTATTTGGAGC 830
Qy 1621 CTGCTGCCAAACCACTTGTGAGTCTTCAAGTCTTCAAGTGAAGATGTCGCTCATCTG 1680
Db 829 CTGCTGCCAAACCACTTGTGAGTCTTCAAGTCTTCAAGTGAAGATGTCGCTCATCTG 770
Qy 1681 GCGTCTGTCACATGGCC-TTTGGGT-GTTCCTCGAGTCTTCAACACAGTGCACCTTG 1738
Db 769 GCGTCTGTCACATGGCC-TTTGGGT-GTTCCTCGAGTCTTCAACACAGTGCACCTTG 710
Qy 1739 GCCAGAGGCGACCGGCTGCTGCGAGACGCTGCGGAGCTCACCTTCTGCTGGGACTCT 1798
Db 709 GCCAGAGGCGACCGGCTGCTGCGAGACGCTGCGGAGCTCACCTTCTGCTGGGACTCT 650
Qy 1799 TCGGTTACTGCTGCTTCTAGTCAATCAAGTGGCTGTGTGTGCTGGGCTGCCAGGCGC 1858
Db 649 TCGGTTACTGCTGCTTCTAGTCAATCAAGTGGCTGTGTGTGCTGGGCTGCCAGGCGC 590
Qy 1859 CCTCG---CCAGAGTCTCATTCACCTTCAACATGTTCTTCTTCTCCACAGACCCCA 1915
Db 589 CCTCGGCGCCAGAGTCTCATTCACCTTCAACATGTTCTTCTTCTCCACAGACCCCA 530
Qy 1916 GCAACAGGCTGCTTACCCCGGCGAGGAGTGTTCAGGCGCACCTGCTGTGCTCTGGCT 1975
Db 529 GCAACAGGCTGCTTACCCCGGCGAGGAGTGTTCAGGCGCACCTGCTGTGCTCTGGCT 470
Qy 1976 TGCGCATGTGCGCATCTCTGCTTGGCAACCCCTGCACTGCTGCAACCGCCACCGCC 2035
Db 469 TGCGCATGTGCGCATCTCTGCTTGGCAACCCCTGCACTGCTGCAACCGCCACCGCC 410
Qy 2036 GCGGCTTGGGAGGCGCGCTGACCGACAGGAGGAGAAACAAGGCGGGTGTCTGAGCC 2095

Db 409 GCCTGCGGAGGAGCCGCTGACGACGAGGAGAAACAAAGCGCGGTTGCTGGACC 350
QY 2096 TGCTCAGCCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGGACAGGGGCTGGATG 2155
Db 349 TGCTCAGCCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGGACAGGGGCTGGATG 290
QY 2156 ATGAAGAGGAGGCGGAGCTCGTCCCTCCGAGGTGCTCATGACACGAGGCGCATCCACCA 2215
Db 289 ATGAAGAGGAGGCGGAGCTCGTCCCTCCGAGGTGCTCATGACACGAGGCGCATCCACCA 230
QY 2216 TCGAGTTCTGCTGGGCTGGCTTCCAAACACCGCTTCTACCTGCGGCTGTGGGCGCTGA 2275
Db 229 TCGAGTTCTGCTGGGCTGGCTTCCAAACACCGCTTCTACCTGCGGCTGTGGGCGCTGA 170
QY 2276 GCCTGGCCACGCCC 2290
Db 169 GCCTGGCCACGCCC 155

RESULT 5

US-11-266-748A-221798
; Sequence 221798, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 221798
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-221798

Query Match 28.8%; Score 760; DB 7; Length 1000;
Best Local Similarity 100.0%; Pred. No. 2.6e-208;
Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1864 CCAGCATCTCATCCATCTCATCAATGTTCTCTCCACAGCCCGCCAGCAACAGG 1923
Db 218 CCAGCATCTCATCCATCTCATCAATGTTCTCTCCACAGCCCGCCAGCAACAGG 277
QY 1924 CTGCTTACCCCGGAGAGGTGTCAGGCCACGCTGGTGGTCTCGCCCTGGCCATG 1983
Db 278 CTGCTTACCCCGGAGAGGTGTCAGGCCACGCTGGTGGTCTCGCCCTGGCCATG 337
QY 1984 GTGCCCATCTCTGCTGGCACACCCCTGACCTGCTGCACCGCCACCGCCGCGCTG 2043
Db 338 GTGCCCATCTCTGCTGGCACACCCCTGACCTGCTGCACCGCCACCGCCGCGCTG 397
QY 2044 CCGAGAGGCGCCGCTGACCGACGAGGAGAAACAAAGGCGCGGTTGCTGGACCTGCTGCAC 2103

Db 398 CCGAGAGGCGCCGCTGACCGACGAGGAGAAACAAAGGCGCGGTTGCTGGACCTGCTGCAC 457
QY 2104 GCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGGACAGGGGCTCGATGATGAAGAG 2163
Db 458 GCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGGACAGGGGCTCGATGATGAAGAG 517
QY 2164 GAGGCGGAGCTCGTCCCTCCGAGGTGCTCATGACACGAGGCGCATCCACACCATCGAGTTC 2223
Db 518 GAGGCGGAGCTCGTCCCTCCGAGGTGCTCATGACACGAGGCGCATCCACACCATCGAGTTC 577
QY 2224 TGCCTGGGCTGGCTTCCAAACACCGCTTCTACCTGCGGCTGTGGGCGCTTGAGCCTGGCC 2283
Db 578 TGCCTGGGCTGGCTTCCAAACACCGCTTCTACCTGCGGCTGTGGGCGCTTGAGCCTGGCC 637
QY 2284 CACGCGGAGCTGTGCGAGGTCTGTGGGCGCATGGTGTGATGCGCATAGGCTTGGGCGCTGGGC 2343
Db 638 CACGCGGAGCTGTGCGAGGTCTGTGGGCGCATGGTGTGATGCGCATAGGCTTGGGCGCTGGGC 697
QY 2344 CCGGAGGTGGGCGTGGGCGCTGTGGTGTGCTGCTCCCATCTTTTGGCGGCTTTGGCGGTGATG 2403
Db 698 CCGGAGGTGGGCGTGGGCGCTGTGGTGTGCTGCTGCTCCCATCTTTTGGCGGCTTTGGCGGTGATG 757
QY 2404 ACCGTGGCTATCTGCTGTGATGAGGAGACTCTCAGCCCTTCTGCAACGCGCTTGGCGGCTG 2463
Db 758 ACCGTGGCTATCTGCTGTGATGAGGAGACTCTCAGCCCTTCTGCAACGCGCTTGGCGGCTG 817
QY 2464 CACTGGGTGGAATTCAGAAACAAAGTTCTACTCAGGCAACGCGCTTACAAAGTGTAGTCCCTTC 2523
Db 818 CACTGGGTGGAATTCAGAAACAAAGTTCTACTCAGGCAACGCGCTTACAAAGTGTAGTCCCTTC 877
QY 2524 ACCTTGGCTGCCACAGATGACTAGGGCCACTGCAAGTCTTCTGCCAGACCTTCTTCTGAC 2583
Db 878 ACCTTGGCTGCCACAGATGACTAGGGCCACTGCAAGTCTTCTGCCAGACCTTCTTCTGAC 937
QY 2584 CTCTGAGGCGAGGAGGAAATAAGACGCTCCGCGCTTGGCA 2623
Db 938 CTCTGAGGCGAGGAGGAAATAAGACGCTCCGCGCTTGGCA 977

RESULT 6

US-11-266-748A-287153
; Sequence 287153, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 287153
; LENGTH: 1000


```
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-287153

Query Match      28.8%; Score 760; DB 7; Length 1000;
Best Local Similarity 100.0%; Pred. No. 2.6e-208;
Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1864 CCAGCATCTCTATCCATCTTCAATCAATGTTCTTCTTCCACAGCCCAACAGG 1923
Db 218 CCCAGCATCTCTATCCATCTTCAATCAATGTTCTTCTTCCACAGCCCAACAGG 277

QY 1924 CTGCTCTACCCCGGAGGAGGTGTCAGGCGCACTGTGTCTTGGCTTGGCCATG 1983
Db 278 CTGCTCTACCCCGGAGGAGGTGTCAGGCGCACTGTGTCTTGGCTTGGCCATG 337

QY 1984 GTGCCATCTCTGTCTGTGGCACACCCCTGTGACCTGTGACACCCAGCCGCGCTG 2043
Db 338 GTGCCATCTCTGTCTGTGGCACACCCCTGTGACCTGTGACACCCAGCCGCGCTG 397

QY 2044 CGAGGAGGCGCGTGTACCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2103
Db 398 CGAGGAGGCGCGTGTACCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 457

QY 2104 GCATCTGTGAATGGCTGGAGCTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 2163
Db 458 GCATCTGTGAATGGCTGGAGCTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 517

QY 2164 GAGGCGGAGCTCGTCCCTCCGAGGTGCTCATGACAGGAGGAGGAGGAGGAGGAG 2223
Db 518 GAGGCGGAGCTCGTCCCTCCGAGGTGCTCATGACAGGAGGAGGAGGAGGAGGAG 577

QY 2224 TGCTGTGGCTGTCTTCAACACCGCTCTTACCTGTGCGCTGTGGGCGCTGAGCCTG 2283
Db 578 TGCTGTGGCTGTCTTCAACACCGCTCTTACCTGTGCGCTGTGGGCGCTGAGCCTG 637

QY 2284 CAGCCCGAGCTGCGAGGTCTGTGGGCGATGATGAGGAGGAGGAGGAGGAGGAGG 2343
Db 638 CAGCCCGAGCTGCGAGGTCTGTGGGCGATGATGAGGAGGAGGAGGAGGAGGAGG 697

QY 2344 CGGAGGTGGGCGTGGCGGCTGTGGTGTCTGCTCCCATCTTTGCGGCTTTGCGGTG 2403
Db 698 CGGAGGTGGGCGTGGCGGCTGTGGTGTCTGCTCCCATCTTTGCGGCTTTGCGGTG 757

QY 2404 ACCGTGGCTATCTGTGTGTGATGAGGAGGAGTCTCAGCCTTCTTGCACGCGCTG 2463
Db 758 ACCGTGGCTATCTGTGTGTGATGAGGAGGAGTCTCAGCCTTCTTGCACGCGCTG 817

QY 2464 CACTGGGTGGAATTCAGAACAGATTTCTACTCAGGCGAGGCTACAGCTGAGTCCCTT 2523
Db 818 CACTGGGTGGAATTCAGAACAGATTTCTACTCAGGCGAGGCTACAGCTGAGTCCCTT 877

QY 2524 ACCTTCGCTGCCACAGATGACTAGGCGCCACTGAGGTCTTGCAGACCTCTTCTGAC 2583
Db 878 ACCTTCGCTGCCACAGATGACTAGGCGCCACTGAGGTCTTGCAGACCTCTTCTGAC 937

QY 2584 CTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2623
Db 938 CTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 977
```

RESULT 7
US-11-266-748A-338582/c
; Sequence 338582, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266, 748A

```
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 338582
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-338582

Query Match      28.8%; Score 760; DB 7; Length 1000;
Best Local Similarity 100.0%; Pred. No. 2.6e-208;
Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1864 CCAGCATCTCTATCCATCTTCAATCAATGTTCTTCTTCCACAGCCCAACAGG 1923
Db 783 CCAGCATCTCTATCCATCTTCAATCAATGTTCTTCTTCCACAGCCCAACAGG 724

QY 1924 CTGCTCTACCCCGGAGGAGGTGTCAGGCGCACTGTGTCTTGGCTTGGCCATG 1983
Db 723 CTGCTCTACCCCGGAGGAGGTGTCAGGCGCACTGTGTCTTGGCTTGGCCATG 664

QY 1984 GTGCCATCTCTGTCTTGGCACACCCCTGTGACCTGTGACCGCCACCGCGCGCTG 2043
Db 663 GTGCCATCTCTGTCTTGGCACACCCCTGTGACCTGTGACCGCCACCGCGCGCTG 604

QY 2044 CGAGGAGGCGCGCTGACCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2103
Db 603 CGAGGAGGCGCGCTGACCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 544

QY 2104 GCATCTGTGAATGGCTGGAGCTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 2163
Db 543 GCATCTGTGAATGGCTGGAGCTCCGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 484

QY 2164 GAGGCGGAGCTCTGCTCCCTCCGAGGTGCTCATGACAGGCGCATCCACCATCGAGTTC 2223
Db 483 GAGGCGGAGCTCTGCTCCCTCCGAGGTGCTCATGACAGGCGCATCCACCATCGAGTTC 424

QY 2224 TGCTGGGCTGCTCTTCCACAGCGGCTCTACTGCGGCTTGGGCGCTGAGCCTGGCC 2283
Db 423 TGCTGGGCTGCTCTTCCACAGCGGCTCTACTGCGGCTTGGGCGCTTGGGCGCTG 364

QY 2284 CAGCCCGAGCTGTCCGAGGTCTGTGGGCGCATGATGCGCATAGGCGCTGGGCTGGCC 2343
Db 363 CAGCCCGAGCTGTCCGAGGTCTGTGGGCGCATGATGCGCATAGGCGCTGGGCTGGCC 304

QY 2344 CGGAGGTGGGCGTGGCGGCTGTGGTGTGCTTCCCATCTTTGCGGCTTTGCGGTGATG 2403
Db 303 CGGAGGTGGGCGTGGCGGCTGTGGTGTGCTTCCCATCTTTGCGGCTTTGCGGTGATG 244

QY 2404 ACCGTGGCTATCTGTGTGTGATGAGGAGGAGTCTCAGCCTTCTTGCAGCGCTTGGCGCTG 2463
Db 243 ACCGTGGCTATCTGTGTGTGATGAGGAGGAGTCTCAGCCTTCTTGCAGCGCTTGGCGCTG 184

QY 2464 CACTGGGTGGAATTCAGAACAGATTTCTACTCAGGCGAGGAGGAGGAGGAGGAGGAG 2523
Db 183 CACTGGGTGGAATTCAGAACAGATTTCTACTCAGGCGAGGAGGAGGAGGAGGAGGAG 124
```

QY 2524 ACCTTCGCTGCCACAGATGACTAGGGCCCACTGCAAGTCTCTGCGAGACCTCTCTCTGAC 2583
Db 123 ACCTTCGCTGCCACAGATGACTAGGGCCCACTGCAAGTCTCTGCGAGACCTCTCTCTGAC 64
QY 2584 CTCTGAGGAGGAGAGGAATAAAGACGGTCCGCCCTGGCA 2623
Db 63 CTCTGAGGAGGAGAGGAATAAAGACGGTCCGCCCTGGCA 24

RESULT 8

US-11-266-748A-397768
; Sequence 397768, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266, 748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 397768
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-397768

Query Match 28.8%; Score 760; DB 7; Length 1000;
Best Local Similarity 100.0%; Pred. No. 2.6e-208;
Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1864 CCAGCATCTCATCTCACATTCATCAACATGTTCTCTTCTCCACAGCCCAACAGG 1923
Db 218 CCAGCATCTCATCTCACATTCATCAACATGTTCTCTTCTCCACAGCCCAACAGG 277
QY 1924 CTGCTCTACCCCGGAGGAGTGTGTCAGGCGACGCTGGTGGTCTCTGGCCCTTGGCCATG 1983
Db 278 CTGCTCTACCCCGGAGGAGTGTGTCAGGCGACGCTGGTGGTCTCTGGCCCTTGGCCATG 337
QY 1984 GTGCCCATCTCTGCTGGTGGACACCCCTGCACTGCTGCAACCGCCGCGCCGCTG 2043
Db 338 GTGCCCATCTCTGCTGGTGGACACCCCTGCACTGCTGCAACCGCCGCGCCGCTG 397
QY 2044 CGGAGGAGCCGCTGACCGACAGAGGAGAAACAGAGCCGGTGTGTCGACCTGCTGAC 2103
Db 398 CGGAGGAGCCGCTGACCGACAGAGGAGAAACAGAGCCGGTGTGTCGACCTGCTGAC 457
QY 2104 GCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGGCAAGGCGGCTGGATGATGAAGAG 2163
Db 458 GCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGGCAAGGCGGCTGGATGATGAAGAG 517
QY 2164 GAGGCGGAGCTGCTCCCTCCGAGGTGCTCATGCCACCGCCATCCACACCATCGAGTTC 2223
Db 518 GAGGCGGAGCTGCTCCCTCCGAGGTGCTCATGCCACCGCCATCCACACCATCGAGTTC 577

QY 2224 TSCCTGGGCTGGTCTCCAAACACCGCTTCTACCTGGGCTGTGGGCTGTGGGCTGTGGGCT 2283
Db 578 TSCCTGGGCTGGTCTCCAAACACCGCTTCTACCTGGGCTGTGGGCTGTGGGCTGTGGGCT 637
QY 2284 CACGCCCAAGCTGTCCGAGGTTCTGTGGGCAATGCTGATGCGCATAGGCTGTGGGCTGTGGGCT 2343
Db 638 CACGCCCAAGCTGTCCGAGGTTCTGTGGGCAATGCTGATGCGCATAGGCTGTGGGCTGTGGGCT 697
QY 2344 CGGAGGCTGGGCTGTGGGCTGTGGTGTGTCCTCCATCTTTTGGCGCTTTTGGCGGTGATG 2403
Db 698 CGGAGGCTGGGCTGTGGGCTGTGGTGTGTCCTCCATCTTTTGGCGCTTTTGGCGGTGATG 757
QY 2404 ACCGTGGCTATCTCTGCTGTGATGAGGAGCTCTCAGCTTCTGCAACCGCTTGGGCTG 2463
Db 758 ACCGTGGCTATCTCTGCTGTGATGAGGAGCTCTCAGCTTCTGCAACCGCTTGGGCTG 817
QY 2464 CACTGGGTGGAATTCAGAAACAAAGTTCTACTCAGGCAAGGCTTCTGCAAGCTGAGTCCCTTC 2523
Db 818 CACTGGGTGGAATTCAGAAACAAAGTTCTACTCAGGCAAGGCTTCTGCAAGCTGAGTCCCTTC 877
QY 2524 ACCTTCGCTGCCACAGATGACTAGGGCCCACTGCGAGGCTCTGCCAGACCTCTCTCTGAC 2583
Db 878 ACCTTCGCTGCCACAGATGACTAGGGCCCACTGCGAGGCTCTGCCAGACCTCTCTCTGAC 937
QY 2584 CTCTGAGGAGGAGAGGAATAAAGACGGTCCGCCCTGGCA 2623
Db 938 CTCTGAGGAGGAGAGGAATAAAGACGGTCCGCCCTGGCA 977

RESULT 9

US-11-266-748A-468814/c
; Sequence 468814, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266, 748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 468814
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-468814

Query Match 28.8%; Score 760; DB 7; Length 1000;
Best Local Similarity 100.0%; Pred. No. 2.6e-208;
Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1864 CCAGCATCTCATCTCACATTCATCAACATGTTCTCTTCTCCACAGCCCAACAGG 1923

```
Db 783 CCCAGCATCTCTCATCCACTTTCATCAACATGTTCTCTTCTCCACAGCCCCAGCAACAGG 724
Qy 1924 CTGCTCTACCCCCCGGCGAGAGGTGGTTCAGGCCACAGCTGFGGTGCTTGGCCCTTGGCCATG 1983
Db 723 CTGCTCTACCCCCCGGCGAGAGGTGGTTCAGGCCACAGCTGFGGTGCTTGGCCCTTGGCCATG 664
Qy 1984 GTGCCCATCTGCTGCTTGGCACACCCCTGCACCTGCTGCACCGCCACCGCCGCGCTG 2043
Db 663 GTGCCCATCTGCTGCTTGGCACACCCCTGCACCTGCTGCACCGCCACCGCCGCGCTG 604
Qy 2044 CGGAGGAGGCGCTGACCGACAGGAGGAAAAACAAGGCCGGGTGCTGACCTCTCCCTGCAC 2103
Db 603 CGGAGGAGGCGCTGACCGACAGGAGGAAAAACAAGGCCGGGTGCTGACCTCTCCCTGCAC 544
Qy 2104 GCATCTGTGAATGCTGAGGCTCGATGAGGAAAAAGGAGGCGGCTGGATGATGAAGAG 2163
Db 543 GCATCTGTGAATGCTGAGGCTCGATGAGGAAAAAGGAGGCGGCTGGATGATGAAGAG 484
Qy 2164 GAGGCCAGGCTGCTCCCTCCGAGGTGCTCATGCACGAGGCCATCCACACCATCGAGTTC 2223
Db 483 GAGGCCAGGCTGCTCCCTCCGAGGTGCTCATGCACGAGGCCATCCACACCATCGAGTTC 424
Qy 2224 TGCTGGGCTGCTCTCAACACCGCTCTCTACCTGCGCCTGTGGGCCCTTGAGCCTGGCC 2283
Db 423 TGCTGGGCTGCTCTCAACACCGCTCTCTACCTGCGCCTGTGGGCCCTTGAGCCTGGCC 364
Qy 2284 CAGCCAGCTGCTCGAGGTTCTGTGGGCAATGATGCGCATGAGCCTGGGCTGGCC 2343
Db 363 CAGCCAGCTGCTCGAGGTTCTGTGGGCAATGATGCGCATGAGCCTGGGCTGGCC 304
Qy 2344 CGGAGGTGGGCTGGGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2403
Db 303 CGGAGGTGGGCTGGGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 244
Qy 2404 ACCGTGCTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2463
Db 243 ACCGTGCTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 184
Qy 2464 CACTGGGTGAATTCAGAAACAAGTTCTACTAGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTG 2523
Db 183 CACTGGGTGAATTCAGAAACAAGTTCTACTAGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTG 124
Qy 2524 ACCTCGCTCCACAGATGACTAGGCGCCACTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2583
Db 123 ACCTCGCTCCACAGATGACTAGGCGCCACTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 64
Qy 2584 CTCTGAGGCAAGGAGGAATAAGACGGTCCGCTCGCA 2623
Db 63 CTCTGAGGCAAGGAGGAATAAGACGGTCCGCTCGCA 24

RESULT 10
US-11-266-748A-76627
; Sequence 76627, Application US/11266748A
; Publication No. US2006013463A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 5815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; PRIOR FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
```

```
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76627
; LENGTH: 1042
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (16)..(16)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (21)..(21)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (65)..(65)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (80)..(80)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (155)..(155)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (169)..(170)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (181)..(181)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (186)..(186)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191)..(191)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (205)..(205)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (241)..(241)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (248)..(248)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (251)..(251)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (284)..(284)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312)..(312)
; OTHER INFORMATION: n is a, c, g, or t
```

```
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (461)..(461)
/ OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-76627

Query Match      22.0%; Score 581; DB 7; Length 1042;
Best Local Similarity 100.0%; Pred. No. 2,9e-157;
Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2042 TGGGAGGAGGCGCGCTGACCGACAGGAGAAACAAGCCGGGTTGCTGGACCTGCCTG 2101
Db 462 TGGGAGGAGGCGCGCTGACCGACAGGAGAAACAAGCCGGGTTGCTGGACCTGCCTG 521
QY 2102 AGCATCTCTGAATGCTGGAGTCCGATGAGAAAAGCAGGGGCTTGGATGATGAG 2161
Db 522 AGCATCTCTGAATGCTGGAGTCCGATGAGAAAAGCAGGGGCTTGGATGATGAG 581
QY 2162 AGGAGCCGAGCTCGTCCCTCCGAGGTCTCATGCCAGGCCATCCACACATCGAGT 2221
Db 582 AGGAGCCGAGCTCGTCCCTCCGAGGTCTCATGCCAGGCCATCCACACATCGAGT 641
QY 2222 TCTGCTGGGCTGGCTCTCAACACACCGCTCTCTACCTGGGCTGTGGGCTTGGGCTGG 2281
Db 642 TCTGCTGGGCTGGCTCTCAACACACCGCTCTCTACCTGGGCTGTGGGCTTGGGCTGG 701
QY 2282 CCCAGCCGAGCTGTCGAGGTTCTGTGGGCCATGGTATGCCATAGGCTGGGCTGG 2341
Db 702 CCCAGCCGAGCTGTCGAGGTTCTGTGGGCCATGGTATGCCATAGGCTGGGCTGG 761
QY 2342 GCCGGAGGTGGGCGCTGGGCGCTGTGGTGTGCTGCCATCTTTGGCGCTTTGCCGTGA 2401
Db 762 GCCGGAGGTGGGCGCTGGGCGCTGTGGTGTGCTGCCATCTTTGGCGCTTTGCCGTGA 821
QY 2402 TGACCTGTGCTATCTCTGTGGTGTGATGAGGAGTCTTCAGCCTTCTGACGCGCTTGGCGC 2461
Db 822 TGACCTGTGCTATCTCTGTGGTGTGATGAGGAGTCTTCAGCCTTCTGACGCGCTTGGCGC 881
QY 2462 TGACCTGTGCTATCTCTGTGGTGTGATGAGGAGTCTTCAGCCTTCTGACGCGCTTGGCGC 2521
Db 882 TGACCTGTGCTATCTCTGTGGTGTGATGAGGAGTCTTCAGCCTTCTGACGCGCTTGGCGC 941
QY 2522 TCACCTTCTGCTGCCAGATGACTAGGCGCCACTGCAGTCTTCCAGACCTTCTTCTG 2581
Db 942 TCACCTTCTGCTGCCAGATGACTAGGCGCCACTGCAGTCTTCCAGACCTTCTTCTG 1001
QY 2582 ACCTGTAGGAGGAGAGAGAAATAAGACGGTCCGCGCTGGC 2622
Db 1002 ACCTGTAGGAGGAGAGAGAAATAAGACGGTCCGCGCTGGC 1042

RESULT 11
US-11-266-748A-109487
/ Sequence 109487, Application US/11266748A
/ Publication No. US20060134663A1
/ GENERAL INFORMATION:
/ APPLICANT: Harkin, Paul
/ APPLICANT: Johnston, Patrick
/ APPLICANT: Mulligan, Karl
/ TITLE OF INVENTION: Transcriptome Microarray Technology and
/ TITLE OF INVENTION: Methods of Using the Same
/ FILE REFERENCE: 55815-0102 (319189)
/ CURRENT APPLICATION NUMBER: US/11/266,748A
/ CURRENT FILING DATE: 2005-11-03
/ PRIOR APPLICATION NUMBER: EP 04105479.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105482.6
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105483.4
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105507.0
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105485.9
```

```
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: EP 04105484.2
/ PRIOR FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/662,276
/ PRIOR FILING DATE: 2005-03-14
/ PRIOR APPLICATION NUMBER: US 60/700,293
/ PRIOR FILING DATE: 2005-07-18
/ NUMBER OF SEQ ID NOS: 483996
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 109487
/ LENGTH: 1042
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (16)..(16)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (21)..(21)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (65)..(65)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (80)..(80)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (155)..(155)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (169)..(170)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (181)..(181)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (186)..(186)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (191)..(191)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (205)..(205)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (241)..(241)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (248)..(248)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (251)..(251)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (284)..(284)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (312)..(312)
/ OTHER INFORMATION: n is a, c, g, or t
```

```
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (461)..(461)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-109487

Query Match      22.0%; Score 581; DB 7; Length 1042;
Best Local Similarity 100.0%; Pred. No. 2.9e-157; Indels 0; Gaps 0;
Matches 581; Conservative 0; Mismatches 0;

Qy 2042 TCGCGAGGAGCCCGCTGACCGACGAGGAGGAAACAAAGCCCGGGTGTCTGGACCTGCGCTG 2101
Db 462 TCGCGAGGAGCCCGCTGACCGACGAGGAGGAAACAAAGCCCGGGTGTCTGGACCTGCGCTG 521

Qy 2102 ACGCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGAGAGGGGGCTGGATGATGAAG 2161
Db 522 ACGCATCTGTGAATGGCTGGAGCTCCGATGAGGAAAGAGAGGGGGCTGGATGATGAAG 581

Qy 2162 AGGAGGCGGAGCTGTCCTCCCTCCGAGGTGCTCATGACAGGCGCATCCACACCATCGAGT 2221
Db 582 AGGAGGCGGAGCTGTCCTCCCTCCGAGGTGCTCATGACAGGCGCATCCACACCATCGAGT 641

Qy 2222 TCTGCTGGGCTGGCTCTCCAAACCGCTCTCTACCTGCGCTGTGGGCGCTGAGCGCTGG 2281
Db 642 TCTGCTGGGCTGGCTCTCCAAACCGCTCTCTACCTGCGCTGTGGGCGCTGAGCGCTGG 701

Qy 2282 CCGACGCGGAGCTGCGAGGCTTCTGTGGGCGATGGTGATGCGCATAGGCGCTGGCGCTGG 2341
Db 702 CCGACGCGGAGCTGCGAGGCTTCTGTGGGCGATGGTGATGCGCATAGGCGCTGGCGCTGG 761

Qy 2342 GCGGGAGGTGGGCGTGGCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2401
Db 762 GCGGGAGGTGGGCGTGGCGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 821

Qy 2402 TGACCGTGGCTATCTCTGCTGTGTGTGAGGAGCTCTCAGCTTCTCTGACGCGCTGCGGC 2461
Db 822 TGACCGTGGCTATCTCTGCTGTGTGTGAGGAGCTCTCAGCTTCTCTGACGCGCTGCGGC 881

Qy 2462 TGACCTGGGTGAATTCAGAAACAAAGTTCTACTCAGGACGCGGCTACAAGCTGAGTCCCT 2521
Db 882 TGACCTGGGTGAATTCAGAAACAAAGTTCTACTCAGGACGCGGCTACAAGCTGAGTCCCT 941

Qy 2522 TCACCTTCGCTGCCACAGATGACTAGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2581
Db 942 TCACCTTCGCTGCCACAGATGACTAGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1001

Qy 2582 ACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2622
Db 1002 ACCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1042

RESULT 12
US-11-266-748A-129438/c
; Sequence 129438, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
```

```
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 129438
; LENGTH: 1042
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (582)..(582)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (731)..(731)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (759)..(759)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (792)..(792)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (795)..(795)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (802)..(802)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (838)..(838)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (852)..(852)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (857)..(857)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (862)..(862)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (873)..(874)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (888)..(888)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (963)..(963)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (978)..(978)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1022)..(1022)
; OTHER INFORMATION: n is a, c, g, or t
```


; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 76628
; LENGTH: 738
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (45)..(45)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-76628

Query Match 18.9%; Score 500; DB 7; Length 738;
Best Local Similarity 99.7%; Pred. No. 3.9e-134;
Matches 600; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2021 TGCACCGCCACCGCGCGCTGCGAGAGAGCCCGCTGACCGACAGAGAGAAACAAGG 2080
DB 137 TGCACCGCCACCGCGCGCTGCGAGAGAGCCCGCTGACCGACAGAGAGAAACAAGG 196
QY 2081 CCGGGTGTCTGGACCTCCCTGACGATCTGTGAATGCTGAGCTCCGATGAGGAAAGG 2140
DB 197 CCGGGTGTCTGGACCTCCCTGACGATCTGTGAATGCTGAGCTCCGATGAGGAAAGG 256
QY 2141 CAGGGGCGCTGGATGATGAAGAGAGAGCCGAGCTCGTCCCTCCGAGGTGCTCATGCACC 2200
DB 257 CAGGGGCGCTGGATGATGAAGAGAGAGCCGAGCTCGTCCCTCCGAGGTGCTCATGCACC 316
QY 2201 AGGCCATCCACACCATGAGTTCTGCTGGGCTGCGTCTCAACACCGCTCTCTACCTGC 2260
DB 317 AGGCCATCCACACCATGAGTTCTGCTGGGCTGCGTCTCAACACCGCTCTCTACCTGC 376
QY 2261 GCCTGTGGGCGCTGAGCTGCGCCACGCGCTGCTCGAGGTCTCTGGGCCCATGTGA 2320
DB 377 GCCTGTGGGCGCTGAGCTGCGCCACGCGCTGCTCGAGGTCTCTGGGCCCATGTGA 436
QY 2321 TGGGCATAGGCTTGGGCGCTGGGCGGAGGTGGGCGGTGCTGGTCTGGTCTCCCA 2380
DB 437 TGGGCATAGGCTTGGGCGCTGGGCGGAGGTGGGCGGTGCTGGTCTGGTCTCCCA 496
QY 2381 TCTTTGCGGCTTTGCGGTGATGACCGTGGCTATCTCTGCTGGTGAATGAGGAGACTCTCAG 2440
DB 497 TCTTTGCGGCTTTGCGGTGATGACCGTGGCTATCTCTGCTGGTGAATGAGGAGACTCTCAG 556
QY 2441 CCTTCTGACGCGCTGCGGCTGACCTGGGTGGATTTCCAGAACAGTTCTACTCAGGCA 2500
DB 557 CCTTCTGACGCGCTGCGGCTGACCTGGGTGGATTTCCAGAACAGTTCTACTCAGGCA 616
QY 2501 CCGGCTACAGCTGAGTCCCTTACCTTCCGCTGACAGATGACTAGGGCCCACTGCAGG 2560
DB 617 CCGGCTACAGCTGAGTCCCTTACCTTCCGCTGACAGATGACTAGGGCCCACTGCAGG 676
QY 2561 TCTTGCAGACCTCTCTTCTGA CTTCTGAGGCGAGGAGGAAATAAGACCGTCCGCCCTG 2620
DB 677 TCTTGCAGACCTCTCTTCTGA CTTCTGAGGCGAGGAGGAAATAAGACCGTCCGCCCTG 736
QY 2621 GC 2622

Db 737 GC 738

RESULT 15

US-11-266-748A-109488
; Sequence 109488, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 109488
; LENGTH: 738
; TYPE: DNA
; ORGANISM: Homo Sapiens
; NAME/KEY: misc feature
; LOCATION: (45)..(45)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-109488

Query Match 18.9%; Score 500; DB 7; Length 738;
Best Local Similarity 99.7%; Pred. No. 3.9e-134;
Matches 600; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2021 TGCACCGCCACCGCGCGCTGCGAGAGAGCCCGCTGACCGACAGAGAGAAACAAGG 2080
DB 137 TGCACCGCCACCGCGCGCTGCGAGAGAGCCCGCTGACCGACAGAGAGAAACAAGG 196
QY 2081 CCGGGTGTCTGGACCTCCCTGACGATCTGTGAATGCTGAGCTCCGATGAGGAAAGG 2140
DB 197 CCGGGTGTCTGGACCTCCCTGACGATCTGTGAATGCTGAGCTCCGATGAGGAAAGG 256
QY 2141 CAGGGGCGCTGGATGATGAAGAGAGAGCCGAGCTCGTCCCTCCGAGGTGCTCATGCACC 2200
DB 257 CAGGGGCGCTGGATGATGAAGAGAGAGCCGAGCTCGTCCCTCCGAGGTGCTCATGCACC 316
QY 2201 AGGCCATCCACACCATGAGTTCTGCTGGGCTGCGTCTCAACACCGCTCTCTACCTGC 2260
DB 317 AGGCCATCCACACCATGAGTTCTGCTGGGCTGCGTCTCAACACCGCTCTCTACCTGC 376
QY 2261 GCCTGTGGGCGCTGAGCTGCGCCACGCGCTGCTCCGAGGTCTCTGGGCCCATGTGA 2320
DB 377 GCCTGTGGGCGCTGAGCTGCGCCACGCGCTGCTGGGCGGTCTCTGGGCCCATGTGA 436
QY 2321 TGGGCATAGGCTTGGGCGCTGCGGCGGAGGTGGGCGGTGCTGGTCTGGTCTCCCA 2380
DB 437 TGGGCATAGGCTTGGGCGCTGCGGCGGAGGTGGGCGGTGCTGGTCTGGTCTCCCA 496

| | | | |
|----|------|--|------|
| QY | 2381 | TCCTTCCGCGCTTTGCGGTGATGACCGTGGCTATCCTGTGGTGATGGAGGACTCTCAG | 2440 |
| Db | 497 | TCCTTCCGCGCTTTGCGGTGATGACCGTGGCTATCCTGTGGTGATGGAGGACTCTCAG | 556 |
| QY | 2441 | CCTTCCTGCA CGCCCTGCGGCTGCACTGGGTGGAATTCAGAA CAAAGTTCTACTCAGGCA | 2500 |
| Db | 557 | CCTTCCTGCA CGCCCTGCGGCTGCACTGGGTGGAATTCAGAA CAAAGTTCTACTCAGGCA | 616 |
| QY | 2501 | CGGGCTACAAAGCTGAGTCCCTTCACTTCGCTGCCACAGATGACTAGGGCCCACTGCAGG | 2560 |
| Db | 617 | CGGGCTACAAAGCTGAGTCCCTTCACTTCGCTGCCACAGATGACTAGGGCCCACTGCAGG | 676 |
| QY | 2561 | TCCTGCCAGACCTTCCTTCCTGACCTCTGAGGAGGAGAGGAATAAAGACGGTCCGCCCTG | 2620 |
| Db | 677 | TCCTGCCAGACCTTCCTTCCTGACCTCTGAGGAGGAGAGGAATAAAGACGGTCCGCCCTG | 736 |
| QY | 2621 | GC 2622 | |
| Db | 737 | GC 738 | |

Search completed: June 30, 2006, 06:49:30
Job time : 310 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2006, 12:59:20 ; Search time 198 Seconds
(without alignments)
1898.141 Million cell updates/sec

Title: US-10-783-519-2

Perfect score: 4266

Sequence: 1 MGSMFSESEVALVQLFLPTA.....HWVEFQNKFYSGTGKLSDF 822

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: A_Geneseq_8:*
- 2: Geneseq1980s:*
- 3: Geneseq1990s:*
- 4: Geneseq2000s:*
- 5: Geneseq2001s:*
- 6: Geneseq2002s:*
- 7: Geneseq2003as:*
- 8: Geneseq2003bs:*
- 9: Geneseq2004s:*
- 10: Geneseq2005s:*
- 11: Geneseq2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------------|
| 1 | 4266 | 100.0 | 822 | 2 | AAW41943 |
| 2 | 4266 | 100.0 | 822 | 5 | ABB81811 Human OC- |
| 3 | 4266 | 100.0 | 822 | 8 | ADRA4119 Human 116 |
| 4 | 4266 | 100.0 | 822 | 9 | ADY92606 Human 116 |
| 5 | 4235.5 | 99.3 | 830 | 8 | ADI28834 Human mod |
| 6 | 4235.5 | 99.3 | 830 | 8 | ADP12465 Protein e |
| 7 | 4235.5 | 99.3 | 830 | 8 | ABM81281 Tumour-as |
| 8 | 3996 | 93.7 | 787 | 8 | ADH22516 Human tra |
| 9 | 3982 | 93.3 | 787 | 8 | ABM84820 Human dia |
| 10 | 3131.5 | 73.4 | 614 | 2 | AAW98112 T-cell me |
| 11 | 3131.5 | 73.4 | 614 | 2 | ABM81282 Tumour-as |
| 12 | 3127.5 | 73.3 | 614 | 2 | AAW98113 T-cell me |
| 13 | 3127.5 | 73.3 | 614 | 6 | ABR43098 Human T-c |
| 14 | 2917.5 | 68.4 | 643 | 5 | ABPA1524 Human ova |
| 15 | 2119 | 49.7 | 856 | 8 | ADH09950 Human hos |
| 16 | 2118.5 | 49.7 | 856 | 2 | AAR33281 TsFl enco |
| 17 | 2116 | 49.6 | 856 | 8 | ADH09951 Human hos |
| 18 | 2116 | 49.6 | 856 | 8 | ADI28835 Human mod |
| 19 | 2116 | 49.6 | 856 | 8 | ADP54581 Human PRO |
| 20 | 2116 | 49.6 | 856 | 8 | ADP23897 PRO polyp |
| 21 | 2116 | 49.6 | 856 | 9 | ADY15610 PRO polyp |
| 22 | 2116 | 49.6 | 856 | 9 | ADY19962 PRO polyp |
| 23 | 2074.5 | 48.6 | 839 | 8 | ABM84971 Human dia |

| | | | | | |
|----|--------|------|------|----|--------------------|
| 24 | 2066 | 48.4 | 818 | 6 | ABR62420 Protein G |
| 25 | 1931 | 45.3 | 840 | 5 | ADP69796 Human pol |
| 26 | 1931 | 45.3 | 840 | 8 | ADH09949 Human hos |
| 27 | 1931 | 45.3 | 840 | 8 | ADI28836 Human mod |
| 28 | 1926.5 | 45.2 | 831 | 8 | ADH09952 Human hos |
| 29 | 1926.5 | 45.2 | 831 | 8 | ADI28837 Human mod |
| 30 | 1918.5 | 45.0 | 831 | 5 | ABB08457 Human tum |
| 31 | 1918.5 | 45.0 | 831 | 8 | ADH09948 Human hos |
| 32 | 1918.5 | 45.0 | 831 | 8 | ADH09953 Human hos |
| 33 | 1918.5 | 45.0 | 841 | 4 | AAB60100 Human tra |
| 34 | 1912 | 44.8 | 830 | 8 | ADH09947 Human hos |
| 35 | 1882 | 44.1 | 777 | 4 | AAB48977 Human imm |
| 36 | 1801.5 | 42.2 | 847 | 8 | ADRA4120 Human OC- |
| 37 | 1801.5 | 42.2 | 847 | 9 | ADY92607 116 KD os |
| 38 | 1720 | 40.3 | 855 | 4 | ABB58420 Drosophil |
| 39 | 1720 | 40.3 | 855 | 10 | AEQ02510 Drosophil |
| 40 | 1701 | 39.9 | 834 | 4 | ABB71636 Drosophil |
| 41 | 1701 | 39.9 | 834 | 4 | ABB71569 Drosophil |
| 42 | 1539.5 | 36.1 | 703 | 6 | ABR62416 V_ATPase |
| 43 | 1539.5 | 36.1 | 935 | 8 | ADN23157 Bacterial |
| 44 | 1538 | 36.1 | 814 | 4 | ABB69950 Drosophil |
| 45 | 1501 | 35.2 | 1030 | 8 | ADN23158 Bacterial |

ALIGNMENTS

RESULT 1

AAW41943

ID AAW41943 standard; protein; 822 AA.

XX AAW41943;

AC AAW41943;

XX 02-JUL-1998 (first entry)

XX Human OC-116 kDa protein.

XX Osteoclast; human; OC-116 kDa; screening; cell surface marker; probe.

XX Homo sapiens.

XX WO9803651-AL.

XX 29-JAN-1998.

XX 10-JUL-1997; 97WO-US012569.

XX 19-JUL-1996; 96US-00684932.

XX (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.

XX Stashenko P, Li Y, Wucherpennig AL;

XX WPI; 1998-120776/11.

XX N-PSDB; AAV04529.

XX Osteoclast specific or related DNA sequence - useful as probe to screen

XX genomic DNA or cDNA library, or as osteoclast cell surface marker.

XX Example 8; Page 61-64; 75pp; English.

XX The present sequence represents the human OC-116 kDa protein sequence

XX which is used in an example of the present invention, which describes

XX isolated osteoclast specific or related DNA sequences. The present

XX invention also describes; a DNA construct capable of replicating and

XX optionally expressing, in a host cell, osteoclast specific or related

XX DNA, comprising an osteoclast specific or related DNA sequence and

XX sequences necessary for transforming or transfecting a host cell, and for

XX replicating and optionally expressing an osteoclast specific or related

XX DNA sequence in a host cell; and a cell stably transformed or transfected

XX with the DNA construct. The osteoclast specific or related DNA sequence

XX can be used as a probe to screen a genomic DNA or cDNA library for

XX osteoclast specific or related DNA sequences, or as an osteoclast cell

[illegible]

| RESULT 2 | ID | ASB81811 standard; protein; 822 AA. | Qy | AVYLALNQCSVSTTHKCLIFAEAWCSVRDLPALQELRDSMSBEGSVAVNHRITPCDDMPEPT | 360 |
|---------------------------|----|-------------------------------------|----|--|-----|
| ASB81811 | XX | | Db | AVYLALNQCSVSTTHKCLIFAEAWCSVRDLPALQELRDSMSBEGSVAVNHRITPCDDMPEPT | 360 |
| ASB81811 | XX | | Qy | LIETNPTASFCQIVDRCVGVGRYQEVNPAPTIITFPPLFAVMPGQDVGHLMLFALAM | 420 |
| 16-SEP-2002 (first entry) | DT | | Db | LIETNPTASFCQIVDRCVGVGRYQEVNPAPTIITFPPLFAVMPGQDVGHLMLFALAM | 420 |

| | |
|----|---|
| DE | Human OC-116KDa. |
| XX | |
| KW | Human; osteoclast; gene therapy; aberrant bone resorption; OC-116KDa. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| FN | US6403304-B1. |
| XX | |
| PD | 11-JUN-2002. |
| XX | |
| PF | 19-JUL-1996; 96US-00684932. |
| XX | |
| PR | 06-APR-1993; 93US-00045270. |
| PR | 23-FEB-1995; 95US-00392678. |
| PR | 20-JUL-1995; 95US-0001292P. |
| PR | 22-FEB-1996; 96US-00605378. |
| XX | |
| PA | (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN. |
| XX | |
| PI | Stashenko P, Li Y, Wucherpfennig AL; |
| XX | |
| DR | WPI; 2002-536031/57. |
| DR | N-PSDE; ABN86735. |
| XX | |
| PT | Novel isolated osteoclast-specific or -related DNA sequence, useful for |
| PT | producing gene products useful in the therapeutic treatment or diagnosis |
| PT | of disorders involving aberrant bone resorption. |
| XX | |
| PS | Example 8; Fig 3; 34pp; English. |
| XX | |
| CC | The invention relates to novel human osteoclast-specific or -related cDNA |
| CC | sequences. The sequence represents human osteoclast 116KDa (OC-116KDa). |
| CC | The sequences may have a use in gene therapy. The sequences of the |
| CC | invention are useful in the production of gene products useful in the |
| CC | therapeutic treatment or diagnosis of disorders involving aberrant bone |
| CC | resorption, for generating peptides which are useful for producing |
| CC | antibodies for identifying osteoclast-specific or -related peptides or |
| CC | gene products |
| XX | |
| SO | Sequence 822 AA; |

| Query Match | 100.0% | Score 4266; | DB 5; | Length 822; |
|-----------------------|-----------------|---|-----------|-------------|
| Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| Matches 822; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 1 | MGSFRSEEVALLQFLPTAAATCVSRGELGLVFRDLNASYSAFORRFVWDVWRCEE | 60 | |
| Db | 1 | MGSFRSEEVALLQFLPTAAATCVSRGELGLVFRDLNASYSAFORRFVWDVWRCEE | 60 | |
| Qy | 61 | LEKTFPLQBEVRRAGLVLPSPKGRLLPAPPPRDLRLRTOETERLAQELRDVRGNOQAURA | 120 | |
| Db | 61 | LEKTFPLQBEVRRAGLVLPSPKGRLLPAPPPRDLRLRTOETERLAQELRDVRGNOQAURA | 120 | |
| Qy | 121 | QLHQQLHAAVALRCQHEPQLAAAAHTDCASERTPLQAPGGPHQDLRVNFAVAGVEPHKAP | 180 | |
| Db | 121 | QLHQQLHAAVALRCQHEPQLAAAAHTDCASERTPLQAPGGPHQDLRVNFAVAGVEPHKAP | 180 | |
| Qy | 181 | ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEGIQGIRKITDCGF | 240 | |
| Db | 181 | ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEGIQGIRKITDCGF | 240 | |
| Qy | 241 | CHVFPFLQCEARLGAQLQOQSOELQEVLGETERFLSQVLGRVLQLPPQGVQVHKMK | 300 | |
| Db | 241 | CHVFPFLQCEARLGAQLQOQSOELQEVLGETERFLSQVLGRVLQLPPQGVQVHKMK | 300 | |
| Qy | 301 | AVYIALNQCSTVTHKCLIAEAMCSVRDLPALQBALRDSMEEGVSVAVHRIPCBDMPEPT | 360 | |
| Db | 301 | AVYIALNQCSTVTHKCLIAEAMCSVRDLPALQBALRDSMEEGVSVAVHRIPCBDMPEPT | 360 | |
| Qy | 361 | LIRTRNRTASFGQIVDRYGVGRYQEVNPAPTYITITFPFLFAVMFGDVGHGLMLFLFALAM | 420 | |
| Db | 361 | LIRTRNRTASFGQIVDRYGVGRYQEVNPAPTYITITFPFLFAVMFGDVGHGLMLFLFALAM | 420 | |

DR N-PSDB: ADI28820.
XX GENBANK: 1924145.
PT Identifying a candidate p53 pathway modulating agent for treating e.g.,
PT cancer, comprises contacting an assay system comprising a MP53
PT polypeptide or nucleic acid with a test agent and detecting a test agent-
PT biased activity.
XX
XX
XX Example 2; SEQ ID NO 25; 139pp; English.
XX
CC In the present invention, genetic screens were designed to identify
CC modifiers of the p53 pathway in caenorhabditis elegans, where a
CC homozygous p53 deletion mutant was used. Various specific genes were
CC silenced by RNA inhibition. Genes causing altered phenotypes in the worms
CC were identified as modifiers of the p53 pathway. Human orthologs, denoted
CC modifiers of p52 (MP53), of these modifiers were identified. These
CC include the present sequence, characterised as T-cell immune regulator 1,
CC ATPase, H+ transporting, or lysosomal V0 protein a isoform 3. MP53
CC nucleic acids and polypeptides are attractive drug targets for the
CC treatment of pathologies associated with a defective p53 signalling
CC pathway, such as cancer. Methods for modulating MP53 function and/or the
CC p53 pathway in a mammalian cell involve contacting the cell with an agent
CC that specifically binds a MP53 polypeptide or nucleic acid. The agent may
CC be a small molecular modulator, a nucleic acid modulator or an antibody.
XX
XX Sequence 830 AA;
Query Match 99.3%; Score 4235.5; DB 8; Length 830;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 820; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 MGSMPFSEVALVQLFLPTAAATCTVSRGELGVFRDLNANASVAFQRFVVDVWRCCE 60
DB 1 MGSMPFSEVALVQLFLPTAAATCTVSRGELGVFRDLNANASVAFQRFVVDVWRCCE 60
QY 61 LKFTFTFLOEVRAGLVLPKPPKGRLPAPPDRLRLRQETERLAQELDRVGNQOALRA 120
DB 61 LKFTFTFLOEVRAGLVLPKPPKGRLPAPPDRLRLRQETERLAQELDRVGNQOALRA 120
QY 121 QLHQLQLAAVLRQGEHPQLAAHTDGASERTLLQAPGGPHQDLRVNFVAGAVEPHKAP 180
DB 121 QLHQLQLAAVLRQGEHPQLAAHTDGASERTLLQAPGGPHQDLRVNFVAGAVEPHKAP 180
QY 181 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEOIGOKIRKIDCFH 240
DB 181 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEOIGOKIRKIDCFH 240
QY 241 CHVFPLQOEARELQALQQLQOQSOELOEVLGETERFLSOVLGRVLQLLPPGQVQHKMK 300
DB 241 CHVFPLQOEARELQALQQLQOQSOELOEVLGETERFLSOVLGRVLQLLPPGQVQHKMK 300
QY 301 AVYALALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLDSSMERGVSAVAHRIPCRDMPT 360
DB 301 AVYALALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLDSSMERGVSAVAHRIPCRDMPT 360
QY 361 LINTNFTASQGVDRYGVGRVQEVNAPYTTITPPFLFANVDFGVGHLLMFLPALAM 420
DB 361 LINTNFTASQGVDRYGVGRVQEVNAPYTTITPPFLFANVDFGVGHLLMFLPALAM 420
QY 421 VLAENRPAKAAQNEIWTQFFRGYLLLLMGLFSIYTGFIYNECFSRATSIFFSGWSVAA 480
DB 421 VLAENRPAKAAQNEIWTQFFRGYLLLLMGLFSIYTGFIYNECFSRATSIFFSGWSVAA 480
QY 481 MANQSGWDAFLAQHTMTLTDPNVTGVFLGPPYFGIDPILWLSAANHLSTLNSFKMKMSVI 540
DB 481 MANQSGWDAFLAQHTMTLTDPNVTGVFLGPPYFGIDPILWLSAANHLSTLNSFKMKMSVI 540
QY 541 LGVHHMAGVVLGVFNHVFQGRHRLLETLPELTFLLGLFGYLVFLVIYKVLGVMAARA 600
DB 541 LGVHHMAGVVLGVFNHVFQGRHRLLETLPELTFLLGLFGYLVFLVIYKVLGVMAARA 600
QY 601 AS-PSLIIHFNNFLFSHSPSNRLLYPROEVQATLVLLAMVPIILLGTPLHLHRRH 659
DB 601 AS-PSLIIHFNNFLFSHSPSNRLLYPROEVQATLVLLAMVPIILLGTPLHLHRRH 659

DB 601 ASAPSILIIHFNNFLFSHSPSNRLLYPROEVQATLVLLAMVPIILLGTPLHLHRRH 660
QY 660 RLRLRRPADRQENKAGLLDLPDASVNGWSSDBEKAGLDDEEAEALVPSEVLHQAIHT 719
DB 661 RLRLRRPADRQENKAGLLDLPDASVNGWSSDBEKAGLDDEEAEALVPSEVLHQAIHT 720
QY 720 IEFCLGCVSNATSYLRWLWLSLAHAQLSEVLWAMVRRIGLGLGREVGAAVVLVPIFAAF 779
DB 721 IEFCLGCVSNATSYLRWLWLSLAHAQLSEVLWAMVRRIGLGLGREVGAAVVLVPIFAAF 780
QY 780 AVMTVAILLVMEGLSAPFLHALRLHWVEFQNKFSYGTGYKLSPPF 822
DB 781 AVMTVAILLVMEGLSAPFLHALRLHWVEFQNKFSYGTGYKLSPPF 823
RESULT 6
ADP12465
ID ADP12465 standard; protein; 830 AA.
XX ADP12465;
XX
XX 12-AUG-2004 (first entry)
XX
XX Protein encoded by mRNA of the invention #75.
XX
XX transplant rejection; immune system; rheumatoid arthritis; lupus;
XX inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
XX Homo sapiens.
XX WO2004042346-A2.
XX
XX 21-MAY-2004.
XX
XX 24-APR-2003; 2003WO-US012946.
XX
XX 24-APR-2002; 2002US-00131831.
XX 20-DEC-2002; 2002US-00325899.
XX
XX (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
XX Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
XX Rosenberg S;
XX WPI; 2004-400724/37.
XX
XX Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
XX pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
XX rejection, in an individual, comprises detecting the expression level of
XX the genes.
XX
XX Claim 65; SEQ ID NO 2474; 1762pp; English.
XX
XX The present invention relates to diagnosing or monitoring transplant
XX rejection, e.g. cardiac or kidney transplant rejection, in an individual
XX comprising detecting the expression level of one or more genes. The
XX methods, system and kits are useful in diagnosing or monitoring
XX transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
XX islet, lung, bone marrow or stem cell transplant rejection,
XX xenotransplant rejection or mechanical organ replacement rejection, in an
XX individual. The method is also useful in assessing the immune status of
XX an individual. The methods are also useful in diagnosing and monitoring
XX diseases that involve the immune system, e.g. rheumatoid arthritis,
XX lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
XX viral, bacterial or fungal infection. The present sequence represents a
XX protein that is encoded by the mRNA of the invention.
XX
XX Sequence 830 AA;
Query Match 99.3%; Score 4235.5; DB 8; Length 830;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 820; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 301 AVYALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLDSSMEEGVSAVAHRIPCRDMPPT 360
DB 301 AVYALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLDSSMEEGVSAVAHRIPCRDMPPT 360
QY 361 LIRNRTASFGQIVDYGVRGYEVNPAVPTIITFFPFLFAVMEGDVGHGLMLFLFALAM 420
DB 361 LIRNRTASFGQIVDYGVRGYEVNPAVPTIITFFPFLFAVMEGDVGHGLMLFLFALAM 420
QY 421 VLAENRPAVKAQAQNEIWOTFFRGYLLMLGLFSIYTGFIYNECFSRATSIFFPSGWSVAA 480
DB 421 VLAENRPAVKAQAQNEIWOTFFRGYLLMLGLFSIYTGFIYNECFSRATSIFFPSGWSVAA 480
QY 481 MANQSGWSDAFLAQHTMLTLDPNVTGVLGPIPGIDPIWLSLAANHLSFNLSPKQKMSVI 540
DB 481 MANQSGWSDAFLAQHTMLTLDPNVTGVLGPIPGIDPIWLSLAANHLSFNLSPKQKMSVI 540
QY 541 LGVHVMAFGVVLGVFNHVFHQHRLLELTPELTFLGLFGYLVFLVLYKWLVCWAARA 600
DB 541 LGVHVMAFGVVLGVFNHVFHQHRLLELTPELTFLGLFGYLVFLVLYKWLVCWAARA 600
QY 601 AS-PSILIHFINMFLPSHSNSNLLYPRQEVQVQATLVVLAAMVPIILLGTPLHLHRHR 659
DB 601 ASAPSILIHFINMFLPSHSNSNLLYPRQEVQVQATLVVLAAMVPIILLGTPLHLHRHR 660
QY 660 RLRRRPPADROENKAGLLDLPASVNGWSSDEKAGGLDDEEAELVPSEVLHMQAIHT 719
DB 661 RLRRRPPADROENKAGLLDLPASVNGWSSDEKAGGLDDEEAELVPSEVLHMQAIHT 720
QY 720 IEFCLGCVSNATSYLRWLWALSQAQLSEVLWAMVMRIGLGLGREGVGAVALVPIFAAF 779
DB 721 IEFCLGCVSNATSYLRWLWALSQAQLSEVLWAMVMRIGLGLGREGVGAVALVPIFAAF 780
QY 780 AVMTVAILLVMEGLSAPFLHALRLHWBFQNKFSYGTGYKLSPP 822
DB 781 AVMTVAILLVMEGLSAPFLHALRLHWBFQNKFSYGTGYKLSPP 823

RESULT 8

ADH22516
ID ADH22516 standard; protein; 787 AA.

XX
AC ADH22516;

XX
DT 11-MAR-2004 (first entry)

XX
DE Human transporter & ion channel (TRICH) protein SeqID14.

XX human; transporters and ion channel; TRICH; cell proliferative;
KW arteriosclerosis; cancer; autoimmune/inflammatory; AIDS; asthma;
KW neurological; epilepsy; stroke; developmental; Cushing's syndrome;
KW hypothyroidism; infection; gene therapy; cytostatic; antiinflammatory;
KW immunosuppressive; antiasthmatic; anticonvulsant; neurotropic;
KW neuroprotective; single nucleotide polymorphism; SNP.

XX Homo sapiens.

XX OS WO2003093444-A2.

XX PN 13-NOV-2003.

XX PD 02-MAY-2003; 2003WO-US014026.

XX PF 03-MAY-2002; 2002US-0377435P.

XX PR 03-MAY-2002; 2002US-0377444P.

XX PR 05-JUN-2002; 2002US-0386497P.

XX PR 11-JUN-2002; 2002US-0388180P.

XX PA (INCY-) INCYTE CORP.

XX Baughn MR, Becha SD, Bulloch SA, Chang H, Elliott VS;

PI Emerling BM, Griffin JA, Hafalia AJA, Ison CH, Jackson AA, Jiang X;

PI Jin P, Kable AE, Khare R, Lee SY, Lee S, Mason PM, Marquis JP;

PI Ramkumar J, Richardson TW, Swarnakar A, Tran UK, Chawla NK;

PI Wilson AD;

XX WPI; 2004-022655/02.

DR N-PSDB; ADH22582.

XX New human transporters and ion channels (TRICH), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant TRICH expression e.g. cancer, AIDS, arteriosclerosis, epilepsy,
PT or infections.

XX Claim 1; SEQ ID NO 14; 448pp; English.

XX This invention relates to novel isolated polynucleotides identified as
CC human transporters and ion channels (TRICH), and the encoded polypeptides
CC thereof. Specifically, it describes using these TRICH molecules, as well
CC as agonists, antagonists, antibodies, expression vectors and host cells,
CC in appropriate screening and toxicity assays to assess the effects of
CC exogenous compounds on TRICH expression. The present invention describes
CC TRICH compositions that are useful in the diagnosis, treatment and
CC prevention of various disorders such as cell proliferative (e.g.
CC arteriosclerosis, cancer), autoimmune/inflammatory (e.g. AIDS, asthma),
CC neurological (e.g. epilepsy, stroke) and developmental (e.g. Cushing's
CC syndrome hypothyroidism) and for infections. Accordingly, these TRICH
CC molecules can be used for gene therapy purposes and exhibit various
CC activities such as cytostatic, antiinflammatory, immunosuppressive,
CC antiasthmatic, anticonvulsant, neurotropic and neuroprotective.

CC Furthermore, a microarray is useful in monitoring or measuring protein-
CC protein interactions, drug-target interactions and gene expression
CC profiles. This polypeptide sequence is a human TRICH protein of the
CC invention.

XX SQ Sequence 787 AA;

Query Match 93.7%; Score 3996; DB 8; Length 787;

Best Local Similarity 94.5%; Pred. No. 0;

Matches 778; Conservative 0; Mismatches 1; Indels 44; Gaps 2;

QY 1 MGSNFRSEVALVOLFLPTAAAYTCVSRGELGLVEPRDLNASVSARFPPVVDWRCEE 60
DB 1 MGSNFRSEVALVOLFLPTAAAYTCVSRGELGLVEPRDLNASVSARFPPVVDWRCEE 60
QY 61 LEKTFLOEEVRRAGLVLPKGRLLPAPPPRDLRIQETTERLAQELDRVGRNQQAALRA 120
DB 61 LEKTFLOEEVRRAGLVLPKGRLLPAPPPRDLRIQETTERLAQELDRVGRNQQAALRA 120
QY 121 QLHQLQHLAAVLRQGHPEPQLAAAHDTGASERTPLQAPGPHQDLRVNFVAGAVEPHKAP 180
DB 121 QLHQLQHLAAVLRQGHPEPQLAAAHDTGASERTPLQAPGPHQDLRVNFVAGAVEPHKAP 180
QY 181 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEGIQKIRKITDCFH 240
DB 181 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEGIQKIRKITDCFH 240
QY 241 CHVFPFLOQBEARLQALQQLQOQSOELQVGLGTERFSLQVLRVLQQLPPGQVQVHKMK 300
DB 241 CHVFPFLOQBEARLQALQQLQOQSOELQVGLGTERFSLQVLRVLQQLPPGQVQVHKMK 300
QY 301 AVYALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLDSSMEEGVSAVAHRIPCRDMPPT 360
DB 301 AVYALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLDSSMEEGVSAVAHRIPCRDMPPT 360
QY 361 LIRNRTASFGQIVDYGVRGYEVNPAVPTIITFFPFLFAVMEGDVGHGLMLFLFALAM 420
DB 361 LIRNRTASFGQIVDYGVRGYEVNPAVPTIITFFPFLFAVMEGDVGHGLMLFLFALAM 420
QY 421 VLAENRPAVKAQAQNEIWOTFFRGYLLMLGLFSIYTGFIYNECFSRATSIFFPSGWSVAA 480
DB 421 VLAENRPAVKAQAQNEIWOTFFRGYLLMLGLFSIYTGFIYNECFSRATSIFFPSGWSVAA 480
QY 481 MANQSGWSDAFLAQHTMLTLDPNVTGVLGPIPGIDPIWLSLAANHLSFNLSPKQKMSVI 540
DB 481 MANQSGWSDAFLAQHTMLTLDPNVTGVLGPIPGIDPIWLSLAANHLSFNLSPKQKMSVI 540

QY 541 LGVHMAFGVGLGVNFHFGQHRLLLETLPELTLLGLFGYLVFLVYKWLCVWAARA 600
 Db 541 LGVHMAFGVGLGVNFHFGQHRLLLETLPELTLLGLFGYLVFLVYKWLCVWAARA 600
 QY 601 AS-PSILIHFMFLPSHSPSNRLLYPRQEVQATLVVLALAMVPILLGLTPLLHLLHRH 659
 Db 601 ASAPSIILIHFMFLPSHSPSNRLLYPRQEVQATLVVLALAMVPILLGLTPLLHLLHRH 660
 QY 660 RLRLRRPADROENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAELVPSEVLHQAIHT 719
 Db 661 RLRLRRPADROENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAELVPSEVLHQAIHT 705
 QY 720 IEFCLGCVSNTASYLRMLWLSLAHAQLSEVLWAMVMRIGLGLGREVGVAALVLPFAAF 779
 Db 706 -----EVLWAMVMRIGLGLGREVGVAALVLPFAAF 737
 QY 780 AVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPP 822
 Db 738 AVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPP 780

RESULT 9

ASM84820

ID ASM84820 standard; protein; 787 AA.

AC ASM84820;

DT 18-NOV-2004 (first entry)

DE Human diagnostic and therapeutic pprotein SEQ ID NO:5069.

KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.

OS Homo sapiens.

XX WO2004023973-A2.

XX 25-MAR-2004.

XX 12-SEP-2003; 2003WO-US028227.

XX 12-SEP-2002; 2002US-0410259P.

XX 12-SEP-2002; 2002US-0410260P.

XX (INCY-) INCYTE CORP.

XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;

XX Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;

XX Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;

XX Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;

XX Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;

XX Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;

XX Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;

XX Patary S, Shi X, Suarez Cu;

XX WPI; 2004-329368/30.

XX N-PSDB; ACN43472.

XX New diagnostic and therapeutic polynucleotides and polypeptides, useful

XX in diagnosing a condition, disease or disorder associated with human

XX molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or

XX in gene mapping.

XX Claim 27; Page: 190pp; English.

XX The invention relates to novel diagnostic and therapeutic polynucleotides

XX selected from one of the 2722 sequences defined in the specification. A

XX polynucleotide of the invention may have a use in gene therapy. The human

XX diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be

XX used to diagnose a particular condition, disease or disorder associated

XX with human molecules, e.g. cell proliferative disorders,

XX autoimmune/inflammatory disorder, developmental disorder, endocrine

XX disorder, neurological disorders, gastrointestinal disorders, or

CC infections caused by virus, bacteria, fungi or parasite. The dithp
 CC molecules may also be used in genetic mapping, in identifying individuals
 CC from minute biological samples, in detecting single nucleotide
 CC polymorphisms, as molecular weight markers, and for somatic or germline
 CC gene therapy. The present sequence represents a dithp protein of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm

XX Sequence 787 AA;

Query Match 93.3%; Score 3982; DB 8; Length 787;

Best Local Similarity 94.4%; Pred. No. 0;

Matches 777; Conservative 0; Mismatches 2; Indels 44; Gaps 2;

QY 1 MGSMFRSEVALVQLPPTAAAYTCVSRIGELGLVEPRDLNASVSFAQRFFVVDVVRCEE 60

Db 1 MGSMFRSEVALVQLPPTAAAYTCVSRIGELGLVEPRDLNASVSFAQRFFVVDVVRCEE 60

QY 61 LEKTFTELQEEVRRAGLVLPKGRIPAPPPDLARIQETERLAQELDRVGNQALRA 120

Db 61 LEKTFTELQEEVRRAGLVLPKGRIPAPPPDLARIQETERLAQELDRVGNQALRA 120

QY 121 QLHQLQLHAAVLRQGHPEQLAAATDASERTPLQAQPGPHQDLRVNFVAGAVEPHKAP 180

Db 121 QLHQLQLHAAVLRQGHPEQLAAATDASERTPLQAQPGPHQDLRVNFVAGAVEPHKAP 180

QY 181 ALERLLWRACROFLIASPRELEQPLEHPVTGSPATMTPLISYWGECIQKIRKITDCPH 240

Db 181 ALERLLWRACROFLIASPRELEQPLEHPVTGSPATMTPLISYWGECIQKIRKITDCPH 240

QY 241 CHVFPLQOEERARLQALQOQSOELOEVLETETERFISOVLGRVLOLLPRQOVQVHKMK 300

Db 241 CHVFPLQOEERARLQALQOQSOELOEVLETETERFISOVLGRVLOLLPRQOVQVHKMK 300

QY 301 AVYLAALNQCSTTHKCLIAEAWCSVRDLPALQEAALRDSMEEGVSVAHAHRIPCRDMPT 360

Db 301 AVYLAALNQCSTTHKCLIAEAWCSVRDLPALQEAALRDSMEEGVSVAHAHRIPCRDMPT 360

QY 361 LIITNRFTASFGIVDYGVRGYQEVNPAPYTIITFPPELFAVMPGDVGHLMLFALAM 420

Db 361 LIITNRFTASFGIVDYGVRGYQEVNPAPYTIITFPPELFAVMPGDVGHLMLFALAM 420

QY 421 VLAENRPVAKAONEIMOTFFRGRYLLMLLGLFSYITGTIYNECFSRATSIIPSGMSVAA 480

Db 421 VLAENRPVAKAONEIMOTFFRGRYLLMLLGLFSYITGTIYNECFSRATSIIPSGMSVAA 480

QY 481 MANQSGWSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWSLAANHLSPFNSFKMMSVI 540

Db 481 MANQSGWSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWSLAANHLSPFNSFKMMSVI 540

QY 541 LGVHMAFGVGLGVNFHFGQHRLLLETLPELTLLGLFGYLVFLVYKWLCVWAARA 600

Db 541 LGVHMAFGVGLGVNFHFGQHRLLLETLPELTLLGLFGYLVFLVYKWLCVWAARA 600

QY 601 AS-PSILIHFMFLPSHSPSNRLLYPRQEVQATLVVLALAMVPILLGLTPLLHLLHRH 659

Db 601 ASAPSIILIHFMFLPSHSPSNRLLYPRQEVQATLVVLALAMVPILLGLTPLLHLLHRH 660

QY 660 RLRLRRPADROENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAELVPSEVLHQAIHT 719

Db 661 RLRLRRPADROENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAELVPSEVLHQAIHT 705

QY 720 IEFCLGCVSNTASYLRMLWLSLAHAQLSEVLWAMVMRIGLGLGREVGVAALVLPFAAF 779

Db 706 -----EVLWAMVMRIGLGLGREVGVAALVLPFAAF 737

QY 780 AVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPP 822

Db 738 AVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPP 780

RESULT 10

AAW98112
 ID AAW98112 standard; protein; 614 AA.
 AC AAW98112;
 XX
 DT 21-JUN-1999 (first entry)
 XX
 DE T-cell membrane protein TIRC7.
 XX
 KW TIRC7; T-cell immune response cDNA 7; T-cell membrane protein; human;
 KW immunosuppressive; organ transplantation; graft rejection;
 KW transplant rejection; autoimmune disease; allergy; infection; tumour;
 KW cancer; rheumatoid arthritis; lupus erythematosus; multiple sclerosis;
 KW encephalitis; vasculitis; diabetes mellitus; pancreatitis; gastritis;
 KW thyroiditis; asthma; lepramatosus; therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Protein 1..601
 FT Domain /note= "separately claimed in Claim 1"
 FT Domain 173..209
 FT Domain /note= "transmembrane domain"
 FT Domain 229..246
 FT Domain /note= "transmembrane domain"
 FT Domain 319..346
 FT Domain /note= "transmembrane domain"
 FT Domain 356..400
 FT Domain /note= "transmembrane domain"
 FT Domain 413..437
 FT Domain /note= "transmembrane domain"
 FT Domain 513..532
 FT Domain /note= "transmembrane domain"
 FT Domain 544..585
 FT Domain /note= "transmembrane domain"
 XX
 WO9911782-A1.
 XX
 PD 11-MAR-1999.
 XX
 PF 28-AUG-1998; 98WO-EP005462.
 XX
 PR 29-AUG-1997; 97DE-01038710.
 PR 12-FEB-1998; 98DE-02002653.
 XX
 PA (BGM) BRIGHAM & WOMENS HOSPITAL.
 PA (UTKU/) UTKU N.
 XX
 PI Utku N, Gullans SR, Milford EL;
 XX
 DR WPI; 1999-205186/17.
 DR N-PSDB; AAX24912.
 XX
 PT A novel T-cell transmembrane protein (TIRC7) - useful for modulating
 PT immune responses, for use in organ transplantations and treatment of
 PT infectious disease.
 XX
 PS Claim 1(i); Page 84-86; 97pp; English.
 XX
 CC The present sequence is TIRC7, a novel T-cell transmembrane protein that
 CC exhibits a central role in T-cell activation in vitro and in vivo. TIRC7
 CC mRNA is transiently up-regulated in the early phase of T-cell activation
 CC via a calcineurin-dependent pathway. TIRC7 cDNA (see AAX24912) was
 CC identified following a screen for genes that are differentially expressed
 CC in alloactivated human T cells. A second TIRC7 protein (see AAW98113) is
 CC identical except for one amino acid substitution (Arg121 to Gln),
 CC suggesting 2 alleles of the TIRC7 gene. The invention provides vectors,
 CC host cells, antibodies, TIRC7 peptides, polypeptides and polynucleotides
 CC that are useful in diagnostic compositions for identifying T-cell
 CC activating or co-stimulating compounds, or for identifying inhibitors of
 CC T-cell activation and stimulation. TIRC7 peptides or polypeptides,
 CC vectors and antibodies are useful in the preparation of pharmaceutical
 CC compositions for treatment of acute and chronic diseases involving T-cell

CC activation and Th1 and Th2 immune response, for the treatment of acute
 CC and chronic rejection of allo- and xeno organ transplants and bone marrow
 CC transplantation, for the treatment of rheumatoid arthritis, lupus
 CC erythematosus, multiple sclerosis, encephalitis, vasculitis, diabetes
 CC mellitus, pancreatitis, gastritis, thyroiditis, for the treatment of
 CC malign disorders of T, B or NK cells, for the treatment of asthma,
 CC lepramatosus, Helicobacter pylori associated gastritis, or for the
 CC treatment of skin, adrenal or lung tumors (claimed)
 XX
 SQ Sequence 614 AA;
 Query Match 73.4%; Score 3131.5; DB 2; Length 614;
 Best Local Similarity 99.7%; Pred. No. 7.1e-301;
 Matches 605; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 217 MTFLLSYWGEIQIGQIKIRKIDTCFCHVFPLQOEAEALGALQQLQQQSQELVGLGTER 276
 DB 1 MTFLLSYWGEIQIGQIKIRKIDTCFCHVFPLQOEAEALGALQQLQQQSQELVGLGTER 60
 QY 277 FLSQVLGRVLQQLPPGQVQVHKMAYLALNQSVSTTHKCLIAEAWCSVRDLPALQEAL 336
 DB 61 FLSQVLGRVLQQLPPGQVQVHKMAYLALNQSVSTTHKCLIAEAWCSVRDLPALQEAL 120
 QY 337 RDSMEEGVSAVAHRIPCRDMPPTLIRTNFTASFOGIVDQRYGVGRYQVNPAPYTIITF 396
 DB 121 RDSMEEGVSAVAHRIPCRDMPPTLIRTNFTASFOGIVDQRYGVGRYQVNPAPYTIITF 180
 QY 397 PFLFAVMFGDVGHGLLMFLFALAMVLAENRPAVKAQNEIWQTFRRGRYLLMLGLFSIY 456
 DB 181 PFLFAVMFGDVGHGLLMFLFALAMVLAENRPAVKAQNEIWQTFRRGRYLLMLGLFSIY 240
 QY 457 TGFYINCEFSRATSIFFSGWSVAAMANQSGWSDAFLAQHTMLTLDPNVTGVFLGPYPFGI 516
 DB 241 TGFYINCEFSRATSIFFSGWSVAAMANQSGWSDAFLAQHTMLTLDPNVTGVFLGPYPFGI 300
 QY 517 DPIWLSLAANHLSPLNSFKMKMSVILGVVHMAFGVILGVFNHVFQGRHRLLETLPLTLF 576
 DB 301 DPIWLSLAANHLSPLNSFKMKMSVILGVVHMAFGVILGVFNHVFQGRHRLLETLPLTLF 360
 QY 577 LLGLFGYLVFLVIYKWLVCVMAARAAS-PSILIHFINMFLPSHSPSNRLLYPROEVVQATL 635
 DB 361 LLGLFGYLVFLVIYKWLVCVMAARAASPSILIHFINMFLPSHSPSNRLLYPROEVVQATL 420
 QY 636 VVLALAMVPILLGTPLHLHRRRLRRRPPADQENKAGLLDLPDASVNGWSSDSEKA 695
 DB 421 VVLALAMVPILLGTPLHLHRRRLRRRPPADQENKAGLLDLPDASVNGWSSDSEKA 480
 QY 696 GGLDDEEAELVPSEVLMEHQAHTIEFCIGCVSNTASYLRLWALSALAHQALSEVLWAMVM 755
 DB 481 GGLDDEEAELVPSEVLMEHQAHTIEFCIGCVSNTASYLRLWALSALAHQALSEVLWAMVM 540
 QY 756 RIGLGLGREVGVAAVLVLPFAAFVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYSGT 815
 DB 541 RIGLGLGREVGVAAVLVLPFAAFVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYSGT 600
 QY 816 GYKLSPP 822
 DB 601 GYKLSPP 607
 RESULT 11
 ABM81282
 ID ABM81282 standard; protein; 614 AA.
 XX
 AC ABM81282;
 XX
 DT 18-NOV-2004 (first entry)
 XX
 DE Tumour-associated antigenic target (TAT) polypeptide PRO24831, SEQ:3307.
 DE Tumour-associated antigenic target; TAT; human; overexpression; cancer;
 KW tumour; diagnosis; cell proliferative disorder; breast cancer;
 KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;

central nervous system cancer; bladder cancer; pancreatic cancer;
cervical cancer; melanoma; leukaemia; hybridisation probe;
chromosome identification; chromosome mapping; gene mapping;
gene therapy; cytostatic.
Homo sapiens.
WO2004030615-A2.
15-APR-2004.
29-SEP-2003; 2003WO-US028547.
02-OCT-2002; 2002US-0414971P.
(GETH) GENENTECH INC.
Wu TD, Zhang Z, Zhou Y;
WPI; 2004-347921/32.
N-PSDB; ACN39268.
New tumor-associated antigenic target polypeptides and nucleic acids,
useful in preparing a medicament for treating or detecting a
proliferative disorder, e.g. breast, lung, colorectal, ovarian or
prostate cancer or tumor.
Claim 12; SEQ ID NO 3307; 7273pp; English.
The invention relates to human tumor-associated antigenic target (TAT)
polypeptides, and their related nucleic acids. The TAT polypeptides are
overexpressed in cancer tissues compared to normal tissues, and may thus
serve as effective targets for the diagnosis and treatment of cancer in
mammals. The invention also relates to nucleic acid and polypeptide
sequences at least 80% identical to the TAT nucleic acids and
polypeptides; expression vectors and host cells comprising a TAT nucleic
acid; an antibody specific for a TAT polypeptide; a peptide or organic
molecule which binds to a TAT polypeptide; fusion proteins comprising a
TAT polypeptide; and methods and compositions for the treatment or
diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
antibodies, antagonists, binding molecules and compositions are useful
for diagnosing or treating a cell proliferative disorder associated with
increased TAT expression, particularly cancers such as breast cancer,
colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
cancer, pancreatic cancer, cervical cancer, cancers of the central
nervous system, melanoma and leukemia. TAT nucleic acids may further be
used as hybridisation probes, in chromosome and gene mapping, in
chromosome identification and in gene therapy. The present sequence
represents a TAT polypeptide of the invention
Sequence 614 AA;
Query Match 73.4%; Score 3131.5; DB 8; Length 614;
Best Local Similarity 99.7%; Pred. No. 7.1e-301;
Matches 605; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
217 MTFLSYWGEQIGQKIRKTDCHVHPFLQEEARLQALQLOQSQELQEVIGETER 276
1 MTFLSYWGEQIGQKIRKTDCHVHPFLQEEARLQALQLOQSQELQEVIGETER 60
277 FLSQVLGRVQLQLPPGQVQVHKMAVYALANQCSVSTTHKCLIAEAWCSVRDLPALQAL 336
61 FLSQVLGRVQLQLPPGQVQVHKMAVYALANQCSVSTTHKCLIAEAWCSVRDLPALQAL 120
337 RDSSNEEGVSVAHRIPCEDMPTLIRNRTFASFGIVDRYGVGRYQEVNPAPTYITTF 396
121 RDSSNEEGVSVAHRIPCEDMPTLIRNRTFASFGIVDRYGVGRYQEVNPAPTYITTF 180
397 PFLFVAMFGDVGHLLMFLFALAMVLAENRPAVKAQNEIWTFFRGRVLLLLMGLFSY 456
181 PFLFVAMFGDVGHLLMFLFALAMVLAENRPAVKAQNEIWTFFRGRVLLLLMGLFSY 240
457 TGFYINECFSRATSPFGSGWSVAAMNOSGWSDAFLAQTMLTLDPNVTGVFLGPYPFGI 516

241 TGFYINECFSRATSPFGSGWSVAAMNOSGWSDAFLAQTMLTLDPNVTGVFLGPYPFGI 300
517 DPINSLAANHLSFLNSFKMKMSVILGVVHMAPGVVGVNHHVFGQRHLLLETLPETTF 576
301 DPINSLAANHLSFLNSFKMKMSVILGVVHMAPGVVGVNHHVFGQRHLLLETLPETTF 360
577 LGLFGYLVFLVYKWLVCVMAARAAS-PSILHFNMFHSPSNRLLYPROEVVQATL 635
361 LGLFGYLVFLVYKWLVCVMAARAASPSILHFNMFHSPSNRLLYPROEVVQATL 420
636 VVLALAMVPILLGLTFLHLLHRRRLRRRPPADROENKAGLLDLPDASVNGWSSDEEKA 695
421 VVLALAMVPILLGLTFLHLLHRRRLRRRPPADROENKAGLLDLPDASVNGWSSDEEKA 480
696 GGLDDEEAEELVPSEVLHMQAIHTTIEPCIGCVSNTASYLRMLWALSIAHQAQLSEVLWAMVM 755
481 GGLDDEEAEELVPSEVLHMQAIHTTIEPCIGCVSNTASYLRMLWALSIAHQAQLSEVLWAMVM 540
756 RIGLGLGREVGVAAVLVPIPAFAFAMTVAILLVMEGLSAFLHALRLHWHVEFQNKPYSGT 815
541 RIGLGLGREVGVAAVLVPIPAFAFAMTVAILLVMEGLSAFLHALRLHWHVEFQNKPYSGT 600
816 GYKLSPPF 822
601 GYKLSPPF 607
RESULT 12
AAW98113
ID AAW98113 standard; protein; 614 AA.
AC AAW98113;
XX 21-JUN-1999 (first entry)
DT T-cell membrane protein TIRC7.
DE
KW TIRC7; T-cell immune response cDNA 7; T-cell membrane protein; human;
immunosuppressive; organ transplantation; graft rejection;
transplant rejection; autoimmune disease; allergy; infection; tumour;
cancer; rheumatoid arthritis; lupus erythematosus; multiple sclerosis;
encephalitis; vasculitis; diabetes mellitus; pancreatitis; gastritis;
thyroiditis; asthma; lepramatosus; therapy.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT Protein 1..601
FT Domain /note= "separately claimed in Claim 1"
FT Domain 173..209
FT Domain /note= "transmembrane domain"
FT Domain 229..246
FT Domain /note= "transmembrane domain"
FT Domain 319..346
FT Domain /note= "transmembrane domain"
FT Domain 356..400
FT Domain /note= "transmembrane domain"
FT Domain 413..437
FT Domain /note= "transmembrane domain"
FT Domain 513..532
FT Domain /note= "transmembrane domain"
FT Domain 544..585
FT Domain /note= "transmembrane domain"
XX WO9811782-A1.
PN
XX 11-MAR-1999.
PD
XX 28-AUG-1998; 98WO-EP005462.
PF
XX 29-AUG-1997; 97DE-01038710.
PR 12-FEB-1998; 98DE-02002653.

XX (BGM) BRIGHAM & WOMENS HOSPITAL.
PA (UTKU/) UTKU N.
XX
XX Utku N, Gullans SR, Milford EL;
XX
XX WPI; 1999-205186/17.
DR N-PSDB; AAX24913.
XX
XX A novel T-cell transmembrane protein (TIRC7) - useful for modulating
PT immune responses, for use in organ transplantations and treatment of
PT infectious disease.
XX
XX Claim 1(i); Page 93-95; 97pp; English.
XX
XX The present sequence is TIRC7, a novel T-cell transmembrane protein that
CC exhibits a central role in T-cell activation in vitro and in vivo. TIRC7
CC mRNA is transiently up-regulated in the early phase of T-cell activation
CC via a calcineurin-dependent pathway. TIRC7 cDNA (see AAX24912) was
CC identified following a screen for genes that are differentially expressed
CC in alloactivated human T cells. A second TIRC7 protein (see AAW98112) is
CC identical except for one amino acid substitution (Gln121 to Arg).
CC suggesting 2 alleles of the TIRC7 gene. The invention provides vectors,
CC host cells, antibodies, TIRC7 peptides, polypeptides and polynucleotides
CC that are useful in diagnostic compositions for identifying T-cell
CC activating or co-stimulating compounds, or for identifying inhibitors of
CC T-cell activation and stimulation. TIRC7 peptides or polypeptides,
CC vectors and antibodies are useful in the preparation of pharmaceutical
CC compositions for treatment of acute and chronic diseases involving T-cell
CC activation and Th1 and Th2 immune response, for the treatment of acute
CC and chronic rejection of allo- and xeno organ transplants and bone marrow
CC transplantation, for the treatment of rheumatoid arthritis, lupus
CC erythematosus, multiple sclerosis, encephalitis, vasculitis, diabetes
CC mellitus, pancreatitis, gastritis, thyroiditis, for the treatment of
CC allergic disorders of T, B or NK cells, for the treatment of asthma,
CC lepromatosis, Helicobacter pylori associated gastritis, or for the
CC treatment of skin, adrenal or lung tumors (Claimed)
XX
XX Sequence 614 AA;
XX
Query Match 73.3%; Score 3127.5; DB 2; Length 614;
Best Local Similarity 99.5%; Pred. No. 1.8e-300;
Matches 604; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
QY 217 MTFLLSYWGEIQGKIRKIDCFCHVFPFLOQEEARLQALQQLQSQSBLQVIGETER 276
DB 1 MTFLLSYWGEIQGKIRKIDCFCHVFPFLOQEEARLQALQQLQSQSBLQVIGETER 60
QY 277 FLSOVLGRVLQLPPGQVQVHKMAYLALNQCSTVTHKCLIAEAWCSVRDLPALQEAL 336
DB 61 FLSOVLGRVLQLPPGQVQVHKMAYLALNQCSTVTHKCLIAEAWCSVRDLPALQEAL 120
QY 337 ROSSMEEGVSAVAHRIPCRDMPPTLIRNFTASFOGIVDRYGVGRVQVNPAPYITITF 396
DB 121 QDSSMEEGVSAVAHRIPCRDMPPTLIRNFTASFOGIVDAYGVGRVQVNPAPYITITF 180
QY 397 PFLFAVWFGVGHGLMFLPALAMVLAENRPAVKAQNEIWQTFGRVYLLLMGLFSYI 456
DB 181 PFLFAVWFGVGHGLMFLPALAMVLAENRPAVKAQNEIWQTFGRVYLLLMGLFSYI 240
QY 457 TGPYINCEFSRATSIIPSGHVSVAAMNQCSTVTHKCLIAEAWCSVRDLPALQEAL 516
DB 241 TGPYINCEFSRATSIIPSGHVSVAAMNQCSTVTHKCLIAEAWCSVRDLPALQEAL 300
QY 517 DPTWLSAANHLSPFNKMSVILGVVHMAFGVFNHFGVGRHRLLETLPELTF 576
DB 301 DPTWLSAANHLSPFNKMSVILGVVHMAFGVFNHFGVGRHRLLETLPELTF 360
QY 577 LGLFGYLVFLVYKWLVCWAARAAS-PSILIHFINMFLFSHSPSNRLLYPRQVQATL 635
DB 361 LGLFGYLVFLVYKWLVCWAARAASPSILIHFINMFLFSHSPSNRLLYPRQVQATL 420
QY 636 VVLALAMVPILLGLTGPLHLHRRRLRRRPAQRQENKAGLLDLPDASVNGSSDEEKA 695

DB 421 VVLALAMVPILLGLTGPLHLHRRRLRRRPAQRQENKAGLLDLPDASVNGSSDEEKA 480
QY 696 GGLDDEEAEELVPSEVLHMQAIHTIEFCIGCVSNTASYLRWALSLSLAHQISEVLWAWM 755
DB 481 GGLDDEEAEELVPSEVLHMQAIHTIEFCIGCVSNTASYLRWALSLSLAHQISEVLWAWM 540
QY 756 RIGLIGREVGVAAVLVLPFAFAFVMTVAILLVMEGLSAFLHALRLHWHVEFQNKFSYGT 815
DB 541 RIGLIGREVGVAAVLVLPFAFAFVMTVAILLVMEGLSAFLHALRLHWHVEFQNKFSYGT 600
QY 816 GYKLSPF 822
DB 601 GYKLSPF 607
RESULT 13
ABR43098
ID ABR43098 standard; protein; 614 AA.
XX
AC ABR43098;
XX
DT 03-JUL-2003 (first entry)
XX
DE Human T-cell immune response cDNA 7 (TIRC7) protein SEQ ID NO:5.
XX
KW Human; human leukocyte antigen; HLA; T-cell immune response cDNA 7;
TIRC7; immune response; immune disease; immunosuppressive; antiallergic;
KW antimicrobial; antibacterial; cytostatic; vulnerary.
XX
OS Homo sapiens.
XX
WO2003025000-A2.
XX
PD 27-MAR-2003.
XX
PF 17-SEP-2002; 2002WO-EP010440.
XX
PR 17-SEP-2001; 2001US-0322895P.
PR 17-SEP-2001; 2001US-0322896P.
XX
PA (UTKU/) UTKU N.
XX
PI Utku N;
XX
DR WPI; 2003-381482/36.
XX
PT Peptide useful for the preparation of a pharmaceutical composition for
PT inhibition of an immune response comprises fragment of the amino acid
PT sequence.
PS Claim 5; Fig 4A; 40pp; English.
XX
XX The present invention describes a peptide (P1) comprising a fragment of
CC the amino acid sequence from human leukocyte associated antigen (HLA)
CC class II alpha 2 chain (see ABR43094) or T cell immune response cDNA7
CC (TIRC7) protein (see ABR43098). Also described is a composition
CC comprising at least one (P1). (P1) has immunosuppressive, antiallergic,
CC antimicrobial, antibacterial, cytostatic and vulnerary activities, and
CC can be used as an inhibitor of the proliferation of peripheral blood
CC mononuclear cells (PBMCs), and as an immune response modulator. (P1) can
CC be used for the preparation of a pharmaceutical composition for the
CC inhibition of an immune response. (P1) can also be used for treating
CC graft versus host disease, autoimmune diseases, allergic diseases,
CC infectious diseases, sepsis and tumours. (P1) can be used for the
CC improvement of wound healing, and for inducing or maintaining immune
CC unresponsiveness. The present sequence represents the human TIRC7
CC protein, which is used in the exemplification of the present invention
XX
SQ Sequence 614 AA;

Query Match 73.3%; Score 3127.5; DB 6; Length 614;
Best Local Similarity 99.5%; Pred. No. 1.8e-300;

Matches 604; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
QY 217 MTFLLSYMGQIQGQIRKTKITDCFHCHVFPFLOQEEARLALGALQOQSOELOEVIGETER 276
Db 1 MTFLLSYMGQIQGQIRKTKITDCFHCHVFPFLOQEEARLALGALQOQSOELOEVIGETER 60
QY 277 FLSQVLGRVLQLPPQGVQVHRKAVYLLALNCQSVSTTHKCLIAEAWCSVRDLPALQEAL 336
Db 61 FLSQVLGRVLQLPPQGVQVHRKAVYLLALNCQSVSTTHKCLIAEAWCSVRDLPALQEAL 120
QY 337 ROSSMEEGVSAVAHRIPCRDMPPTLIRTNRTFASFGIVDRYGVGRYQVNPAPYTIITP 396
Db 121 QDSSMEEGVSAVAHRIPCRDMPPTLIRTNRTFASFGIVDRYGVGRYQVNPAPYTIITP 180
QY 397 PFLFVAMFGDVGHLLMFLFALAMVLAENRPVAKAQNEMOTFERGRVYLLLMGLFSY 456
Db 181 PFLFVAMFGDVGHLLMFLFALAMVLAENRPVAKAQNEMOTFERGRVYLLLMGLFSY 240
QY 457 TGFYINECFSRATSIFFPSGWSVAAMANQSGWSDAFLAQHTMLTLDPNVTGVFLGPYPFGI 516
Db 241 TGFYINECFSRATSIFFPSGWSVAAMANQSGWSDAFLAQHTMLTLDPNVTGVFLGPYPFGI 300
QY 517 DPINSLAANHLFSLSFKMGKSVILGVVHMAFGVVLGVFNHVFHFGQRHLLLETLPETPF 576
Db 301 DPINSLAANHLFSLSFKMGKSVILGVVHMAFGVVLGVFNHVFHFGQRHLLLETLPETPF 360
QY 577 LGLFGYLVFLVIYKWLCVWAARAAS-PSILTHFINMFLFSHPNRLLYPROEVVQATL 635
Db 361 LGLFGYLVFLVIYKWLCVWAARAAS-PSILTHFINMFLFSHPNRLLYPROEVVQATL 420
QY 636 VYLALAMPVILLGTPLHLHRRHRRRRPPADROENKAGLLDLPDASVNGWSSDEEKA 695
Db 421 VYLALAMPVILLGTPLHLHRRHRRRRPPADROENKAGLLDLPDASVNGWSSDEEKA 480
QY 696 GGLDDEERAEVLVPSSEVLMHQAHTTIEPCGCVSNTASYLRULWALSIAHAQLSEVLWAMVM 755
Db 481 GGLDDEERAEVLVPSSEVLMHQAHTTIEPCGCVSNTASYLRULWALSIAHAQLSEVLWAMVM 540
QY 756 RIGLGLGREVGVAADVLPFAFPAVMTVAILLVMEGLSAFLHRLHWHVEFQNFYSCT 815
Db 541 RIGLGLGREVGVAADVLPFAFPAVMTVAILLVMEGLSAFLHRLHWHVEFQNFYSCT 600
QY 816 GYKLSPPF 822
Db 601 GYKLSPPF 607

RESULT 14
ID ABP41524
AC ABP41524; standard; protein; 643 AA.

DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HTAES83, SEQ ID NO:2656.

KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive; chromosome 11q13.4-13.5.

OS Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US018569.
XX 07-JUN-2000; 2000US-0209467P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Birse CE, Rosen CA;
XX WPI; 2002-147878/19.
XX N-PSDB; ABQ54601.
XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
XX cancer), immune disorders, cardiovascular disorders and neurological
XX diseases.
XX Claim 11; SEQ ID NO 2656; 2922pp; English.
XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to
XX recombinant vectors and host cells comprising human ovarian antigen
XX polynucleotides, antibodies against human ovarian antigens, and the use
XX of ovarian antigen polynucleotides and polypeptides in diagnosing,
XX treating, prophylaxis or preventing various ovarian and/or breast-related
XX disorders. Such conditions include ovarian cancer and breast cancer, and
XX metastatic tumours of ovarian or breast origin, reproductive system
XX disorders (e.g., infertility, disorders of pregnancy, anovulation,
XX polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
XX disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
XX shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
XX vaginitis), immune disorders (e.g., congenital and acquired
XX immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
XX blood-related disorders (e.g., anaemia), cardiovascular disorders,
XX respiratory disorders, neurological disorders, gastrointestinal disorders
XX and urinary system disorders. Ovarian antigen polypeptides and
XX polynucleotides may also be used in screening for compounds which
XX modulate ovarian antigen expression or activity. The polynucleotides may
XX further be used for gene therapy, chromosome mapping, in the
XX identification of individuals and in forensic analysis, and the
XX polypeptides may be used as food additives or to prepare antibodies
XX useful in disease diagnosis, drug targeting and phenotyping. The present
XX sequence represents a human ovarian antigen of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 643 AA;

Query Match 68.4%; Score 2917.5; DB 5; Length 643;
Best Local Similarity 99.3%; Pred. No. 1.3e-279;
Matches 559; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 67 FLOEEVRAGLVLPKGRLLPAPPDRDLRLQEEETERLAQELDRVGNQALRAQLHQLQ 126
Db 35 FLOEEVRAGLVLPKGRLLPAPPDRDLRLQEEETERLAQELDRVGNQALRAQLHQLQ 94
QY 127 LHAAVLRQGHBPQLAAHTDASERTPLQAPGGPHQDLRVNFAVAGVEPHKAPALERLL 186
Db 95 LHAAVLRQGHBPQLAAHTDASERTPLQAPGGPHQDLRVNFAVAGVEPHKAPALERLL 154
QY 187 WRACGFLIASPRELEQPLEHPVTGEPATMTFLISYMGQIQGQIRKTKITDCFHCHVFPF 246
Db 155 WRACGFLIASPRELEQPLEHPVTGEPATMTFLISYMGQIQGQIRKTKITDCFHCHVFPF 214
QY 247 LQOEERLGAALQOQSOELOEVIGETERFSLQVLGRVLQLPPQGVQVHRKAVYLLAL 306
Db 215 LQOEERLGAALQOQSOELOEVIGETERFSLQVLGRVLQLPPQGVQVHRKAVYLLAL 274
QY 307 NOCSVSTTHKCLIAEAWCSVRDLPALQEALRDSSMEEGVSAVAHRIPCRDMPPTLIRTNR 366
|||||

Db 275 NQCSVSTTHKCLIAEAWCSVRDLPALQALRDSSMEGSAVAXRIPCRDMPPTLIRTNR 334
QY 367 FTASFOGIVDRYGVGRYQEVNPAPTYITITPPFLPAVNFQDVGHGLMLPALAMVLAENR 426
Db 335 FTASFOGIVDAYGVGRYQEVNPAPTYITITPPFLPAVNFQDVGHGLMLPALAMVLAENR 394
QY 427 PAKAAQNEIWTQFFRCRGRYLLMLGLFSIYTGFIYNECFSRATSIIPSGMSVAAANQSG 486
Db 395 PAKAAQNEIWTQFFRCRGRYLLMLGLFSIYTGFIYNECFSRATSIIPSGMSVAAANQSG 454
QY 487 WSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWSLAANHLSPNSFKMSVILGVVHM 546
Db 455 WSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWSLAANHLSPNSFKMSVILGVVHM 514
QY 547 AFGVILGVFNHVFHQHRRLLLETLPELTLLGLFGVILVLYIKWICVWAARAAS-PSI 605
Db 515 AFGVILGVFNHVFHQHRRLLLETLPELTLLGLFGVILVLYIKWICVWAARAAS-PSI 574
QY 606 LIHFIMFLFSHSPSNRLLYPRQ 628
Db 575 LIHFIMFLFSHSPSNRLLYPRQ 597

RESULT 15

ADH09950

ID ADH09950 standard; protein; 856 AA.

XX ADH09950;

XX 11-MAR-2004 (first entry)

XX Human host factor protein, SEQ ID NO 478.

XX antiviral; host factor; retrovirus; HIV; human.

XX Homo sapiens.

XX WO2003094847-A2.

XX 20-NOV-2003.

XX 07-MAY-2003; 2003WO-US014382.

XX 07-MAY-2002; 2002US-0378711P.

XX (UYEM-) UNIV EMORY.

XX Devine SE;

XX WPI; 2004-011998/01.

XX Identifying an antiviral compound useful for treating HIV comprises
PT exposing a cell that expresses a host factor to a candidate compound to
PT identify an agent that inhibits the expression or activity of the host
PT factor.

PS Claim 13; SEQ ID NO 478; 141pp; English.

XX The invention relates to a novel method for identifying an antiviral
CC compound. The novel method comprises exposing a cell that expresses a
CC host factor to a candidate compound to identify an agent that inhibits
CC the expression or activity of the host factor. The novel method involves
CC identifying an antiviral compound, comprising exposing a first cell that
CC expresses a host factor to a candidate compound, determining whether the
CC candidate compound inhibits the expression or activity of the host factor
CC in the first cell, where a candidate compound that inhibits the
CC expression or activity of the host factor in the first cell is a
CC potential antiviral compound, exposing a second cell to the potential
CC antiviral compound and a retrovirus, and determining whether the compound
CC inhibits the ability of the retrovirus to infect or replicate within the
CC second cell, where a potential antiviral compound that inhibits the
CC ability of the retrovirus to infect the second cell is an antiviral
CC compound. The method is useful in identifying antiviral agents, including

CC those that are effective against retroviruses, such as HIV. This sequence
CC represents a human host factor protein used in the antiviral
CC identification method of the invention.

XX SQ Sequence 856 AA;

Query Match 49.7%; Score 2119; DB 8; Length 856;

Best Local Similarity 50.0%; Pred. No. 3.5e-200;

Matches 433; Conservative 137; Mismatches 226; Indels 70; Gaps 15;

QY 1 MGSFMRSEEEVALQFLPTAAATCVSRCLGELGLVEFRDLNANASVAFQRFVVDVMRCEE 60

Db 1 MGSFMRSEEEVALQFLPTAAATCVSRCLGELGLVEFRDLNANASVAFQRFVVDVMRCEE 60

QY 61 LEKTFTFLOBEVRRAGLVLPKGRPLPAPPRDLRIQEETERLAQELRDVRGNQQALRA 120

Db 61 LERILVYLQEIINRADIPLEGEASPPAPPLKQVLEMQEQLQKLEVELEVTNKEKLRK 120

QY 121 QL-----HQLQLHAAVLRQCHE-----PQLAAAHDTGASERTPLLOAPGPHQDL 165

Db 121 NLLELIBEYTHMLRVTKTFVKRNVEFEPTYEEFPSL-----ESDLSLDYSQMQRIGA 171

QY 166 RVNFVAGAVEPHKAPALERLLWRACRGFLIASFRELEOPLHPTGEPATMTMFLISYWG 225

Db 172 KLGFSVGLNQGVKVEAFKMLWRVCKGYTIVSVAELDESLEDGETGEVIKWYVFLISFWG 231

QY 226 EQIGQKIRKITDCFHCHVFPFLQOBEARLQALQOQSQOELQOELGETERFLSQTGLRV 285

Db 232 EQIGHKVKKICDCHVCHVPYPTAEERREIOGLNTRIQLDLYTLVHKTEQDYLQVLCKA 291

QY 286 LQLLPPGOVQVHKMAVYLALNOCVSSTHKLIAEAWCSVRDLPALQALRDSSMEEG- 344

Db 292 AESVYSRVIQVKKOKAIYHMLNMCSPDVTNKLIAEAWCPEADLQDLRLALEEGRESGA 351

QY 345 -VSAVAHRIPCRDMPTLIRTNFTASFOGIVDRYGVGRYQEVNPAPTYITITPPFLPVM 403

Db 352 TIPSFNWIIPTKETPTPTTRTNKFTGQFQIVDAYGVGSYREVNPAFLTITITPPFLPVM 411

QY 404 FGDVGHGLMLPALAMVLAENRPAVKAQNEIWTQFFRCRGRYLLMLGLFSIYTGFIYNE 463

Db 412 FGDVGHGVNLFALLLVLENHPRNQSQ-EIMRMFFNGRYILLMLGLFSVVTGLIYND 470

QY 464 CFSRATSIIPSGMSVAAANQSG-----WSDAFLAQHTMLTLDPNVTGVFLGYP 513

Db 471 CFSKSYNLFSGSNVSAVMSYSSHPHAKMVKMWNDSVVRHNSILQLDPSIPGVFRGPYP 530

QY 514 FGIDPIWSLAANHLSPNSFKMSVILGVVHMAFGVILGVFNHVFHQHRRLLLETLPE 573

Db 531 LGIDPIWNLATNRLTFLNSFKMSVILGIHMTFGVILGIFNHLFRKKFNLYLSIPE 590

QY 574 LTFLLGLFGVFLVYIKWICVWAARA--SPSILIHFINMFLFSHSPSNRLLYPROEVV 631

Db 591 LLFMLCIGVLIPIFYKWL-VPSAETSRVAPSILIEFINMFLFPASKTSG-LYTQGEYV 648

QY 632 QATLVVALAMPVILLGLTPLLHHRHRR-----LRRPADROEENKAGLLDL 680

Db 649 QRVLLVVTALSVPLVFLGKPLFLLMLHNGRSCFVNRSGYTLIRK--DSEEE----- 698

QY 681 PDASVNGSSDEKAGGLDD--EEEA--ELVPSVILMHOAIHTIEFCLGCVSNSTASVLR 736

Db 699 --VSLGSDQDIEBGNHQVDGCREMACERFNGEILMTQVIHSIEYCLGICISTASVLR 756

QY 737 WALSLAHQAQLSEVLMVAMVRIGLGLRGVGAVALVPIPAAPAVMTVAILLVMGLSFA 796

Db 757 WALSLAHQAQLSDVLMVAMVRGLRVDTTYGV--LLLLPVIALFAVLTIFILLIMEGLSFA 814

QY 797 LHALRLHWVEFQNKFGYGYKLSPP 822

Db 815 LHALRLHWVEFQNKFGYGYKLSPP 840

Search completed: June 29, 2006, 13:02:58
Job time : 203 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2006, 13:03:14 ; Search time 44 Seconds
(without alignments)
1797.504 Million cell updates/sec

Title: US-10-783-519-2
Perfect score: 4266
Sequence: 1 MGSMPRSEEEVALVQLFLPTA.....HWVEFQNFYSGTGKLSPF 822

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

PIR 80:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----------|---------------------------------|
| 1 | 2143.5 | 50.2 | 855 | 2 JH0287 | immune regulatory |
| 2 | 1954 | 45.8 | 838 | 2 A54163 | vacuolar ATPase (E |
| 3 | 1931 | 45.3 | 838 | 2 B38556 | vacuolar proton pu |
| 4 | 1539.5 | 36.1 | 935 | 2 S15795 | vacuolar proton pu |
| 5 | 1491.5 | 35.0 | 865 | 2 T18565 | probable H ⁺ -export |
| 6 | 1406.5 | 33.0 | 873 | 2 T16282 | hypothetical prote |
| 7 | 1375 | 32.2 | 1236 | 2 T19492 | hypothetical prote |
| 8 | 1331 | 31.2 | 821 | 2 H84600 | probable vacuolar |
| 9 | 1323.5 | 31.0 | 843 | 2 T06068 | probable proton pu |
| 10 | 1318 | 30.9 | 840 | 1 A42970 | H ⁺ -exporting ATPas |
| 11 | 1280.5 | 30.0 | 805 | 2 T37787 | probable vacuolar |
| 12 | 1210 | 28.4 | 780 | 2 H84685 | probable vacuolar |
| 13 | 1190 | 27.9 | 775 | 2 T46719 | H ⁺ -exporting ATPas |
| 14 | 1144 | 26.8 | 890 | 2 S54554 | hypothetical prote |
| 15 | 448 | 10.5 | 163 | 2 T46449 | hypothetical prote |
| 16 | 301 | 7.1 | 686 | 2 F72655 | ATP synthase, subu |
| 17 | 283.5 | 6.6 | 658 | 2 E69228 | H ⁺ -transporting tw |
| 18 | 281.5 | 6.6 | 649 | 2 T45102 | probable V-type so |
| 19 | 280.5 | 6.6 | 659 | 2 E71214 | H ⁺ -transporting AT |
| 20 | 273.5 | 6.4 | 659 | 2 A75029 | ATP synthase subun |
| 21 | 270.5 | 6.3 | 701 | 2 E90202 | H ⁺ -transporting AT |
| 22 | 269.5 | 6.3 | 676 | 2 F69394 | H ⁺ -transporting tw |
| 23 | 257.5 | 6.1 | 695 | 2 G64327 | H ⁺ -transporting AT |
| 24 | 255.5 | 6.0 | 722 | 2 F84364 | atpi protein - Ent |
| 25 | 235.5 | 5.5 | 664 | 2 B53610 | probable V-type AT |
| 26 | 203.5 | 4.8 | 622 | 2 B71326 | probable V-type AT |
| 27 | 182 | 4.3 | 454 | 2 F71313 | probable ATP synth |
| 28 | 168.5 | 3.9 | 649 | 2 G71530 | ATP synthase, chai |
| 29 | 167.5 | 3.9 | 649 | 2 C81687 | |

30 163 3.8 690 2 D75487 v-type ATP synthas
31 155.5 3.6 660 2 E81549 ATP synthase, chai
32 155.5 3.6 660 2 C86502 ATP synthase subun
33 154.5 3.6 660 2 E72121 ATP synthase chain
34 141.5 3.3 663 2 C95153 v-type sodium ATP
35 128.5 3.0 378 2 S55634 glycoprotein M - e
36 119 2.8 608 2 C70111 v-type ATPase, sub
37 117.5 2.8 1119 2 AC0045 probable membrane
38 117 2.7 594 2 D95286 hypothetical prote
39 116.5 2.7 704 2 D82227 conserved hypothet
40 113.5 2.7 431 2 AF2635 exopolysaccharide
41 113.5 2.7 435 2 E97417 exoQ-like protein
42 112.5 2.6 1037 2 A36096 Ca2+-transporting
43 112 2.6 669 2 E71127 hypothetical prote
44 111.5 2.6 627 2 B44409 gamma-aminobutyric
45 111.5 2.6 730 2 E75387 NADH dehydrogenase

ALIGNMENTS

RESULT 1

JH0287

immune regulatory protein precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004

C;Accession: JH0287

R;Lee, C.K.; Ghoshal, K.; Beaman, K.D.

Mol. Immunol. 27, 1137-1144, 1990

A;Title: Cloning of a cDNA for a T cell produced molecule with a putative immune regul

A;Reference number: JH0287; MUID:91061805; PMID:2247090

A;Accession: JH0287

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-855 <LEE>

A;Cross-references: UNIPROT:P15920; UNIPARC:UPI00000161FAB; GB:X55184; NID:g52754; PIDN

C;Superfamily: vacuolar ATP synthase 95K chain

Query Match 50.2%; Score 2143.5; DB 2; Length 855;

Best Local Similarity 51.4%; Pred. No. 3.3e-148;

Matches 441; Conservative 129; Mismatches 233; Indels 55; Gaps 15;

Qy 1 MGSMPRSEEEVALVQLFLPTAAYTCVSRGELGLVFRDLNASVSAPQRFVVDVWRCCE 60
Db 1 MGSMPRSEEEVALVQLFLPTAAYTCVSRGELGLVFRDLNASVSAPQRFVVDVWRCCE 60
Qy 61 LEKTFTEFLOEVRERAGLVLPKGRPLAPPDRLLRIQETERLAOELRDVGRNQOALRA 120
Db 61 LEKTFTEFLOEVRERAGLVLPKGRPLAPPDRLLRIQETERLAOELRDVGRNQOALRA 120
Qy 121 QL-----HQLQHAALVRQGH-----PQLAAHTDGASERTPLLAQPGPHQDUVRN 168
Db 121 NLLELVEYTHMLRVTKTFLKRVNVEFTEYEEFPALENDSLDYS-CMQLGA-----KLQ 174
Qy 169 FVAGAVEPKAPALERLLWRACRGLFIASFRELEQLEHPVTGEPATWMTFLSYWGEQI 228
Db 175 FVSGGLIQQRVEAFERMLWRACKGYITVTVAELEDECLDPETGEVINKWYVFLISFWGEQI 234
Qy 229 GQKIRKITDCFHCHVFPPFLQEEARLQALQOQOQELQELVETGERFLSOVLGRVLQ 288
Db 235 GHVKVKICDCYCHCIYIPYNTAERREIQEGLNTRIQDLTVLHKETEDYURQLCKAAES 294
Qy 289 LPPEQVQVHKMAVYALALNQCSVSTTHKCLIAEAWCSVRDLPALQOELRDSSMBEG--VS 346
Db 295 VCSRVVQVRKMAIYHMLNMCSPDVTNKLIAEAWCPEVDLPGLRRALBEGSREGSATIP 354
Qy 347 AVAHRIPCDMPPTLIRTRWTFASFGQIVDRYGVGRYQEVNPAFYITITPPFLFANMFGD 406
Db 355 SPNNTIPTKETPTLIRTNKFTQEGFQNIYDAYGVGSYREVNPAFTIITITPPFLFANMFGD 414
Qy 407 VGHGLMELFALAMVLAENRPAVKAQNETWQTFEGRYLLLMGLPSIYTGFTYNECFPS 466
Db 415 FGHGFVWFLFALLVLNENHPRUSQSQ-EILRMFFDGRYLLLMGLFVSVTGLIYNDGCFPS 473

```
QY 467 RATSIPFSCSVAAANQSG-----WSDAFLAQHTMLTLDPNVTGVFLGPVPGI 516
Db 474 KSVNLFSGSNVCMAYSSSHSBEQBQKMWLWMDSTIRHRTTQLQDNIPGVPRGPYPPI 533
QY 517 DPWLSAANHLFLSFKMKGVIIVGVHMAFGVVLGVFNHVFHFGQRHLLLETLPETLF 576
Db 534 DPWMLATNRLTFLSFKMKGVIILGI PHMTGVVLGIENHLHFRKKNVILSVPEILF 593
QY 577 LIGLFGYLVLYVYKWLVCWAARA-ASPSILHIFINMFLFSPNSRLLYPRQEVQATL 635
Db 594 MLCIFGYLFMIYIKWLAYSAREAPSILIEFINMFLFSPSKTHG-LYPOQAHVQVYL 652
QY 636 VYLALAMPILLGLPLHLHRRRR-----LRRRPADROENKAGLLDLPDAS 684
Db 653 VALTVLAVPLVFLGKPLFLMLHNGRCFMRSGVTLVRK--DSEEE--VSLNGNDIE 708
QY 685 VNGSSDEEKAGLLDDEEBARLPSEVLHMQAHTIEFCLGCVSNTASYLRWLWLSLAHA 744
Db 709 -EGNSRMEB--GCREVTCEEFPGEILLMTQAIHSIEYCLGICISNTASYLRWLWLSLAHA 764
QY 745 QLSEVLWAMVRIGLGLRGVGAUVLPDFAAFVMTVAILLVMEGLSAFLHRLHW 804
Db 765 QLSDLVLAWLKRVGLRVDTTYG---VLLPVMFAFVLTIFILLVMEGLSAFLHRLHW 821
QY 805 VEFQNKFSYGTGYKLSPPF 822
Db 822 VEFQNKFYVAGTKKVPFP 839

RESULT 2
AS4163
vacuolar ATPase (EC 3.6.1.-) 116k chain long form - bovine
C/Species: Bos primigenius taurus (cattle)
C/Date: 07-Jul-1995 #sequence revision 07-Jul-1995 #text_change 09-Jul-2004
C/Accession: AS4163
R.Peng, S.B.; Crider, B.P.; Xie, X.S.; Stone, D.K.
J. Biol. Chem. 269, 17262-17266, 1994
A/Title: Alternative mRNA splicing generates tissue-specific isoforms of 116-kDa polypep
A/Reference number: AS4163; MUID:94274722; PMID:8006034
A/Accession: AS4163
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-838 <PER>
A/Cross-references: UNIPROT:Q29466; UNIPARC:UPI0000138CEA; GB:L31770; NID:g469231; PIDN:
C/Superfamily: vacuolar ATP synthase 95K chain
C/Keywords: alternative splicing; hydrolase

Query Match 45.8%; Score 1954; DB 2; Length 838;
Best Local Similarity 47.7%; Pred. No. 2.2e-134;
Matches 405; Conservative 138; Mismatches 256; Indels 50; Gaps 10;
QY 1 MGSMPRSEVALVOLFLPTAAAYTCVSRIGELGLVEFRDLNASVAFQRRFVVDVVRCEE 60
Db 1 MGELFRSEMTLAQLFLQSEAAAYCCVSELGKGVQFRDLNPDVNVFQKFFVNEVRCEE 60
QY 61 LEKTFPLQEEVRRAGLVLPKGRLPAPPDRDLRIQETERLAQELRDVRGNQOALRA 120
Db 61 MDRKLFVFEKEIRKANIPMDTGENPEVFPFRDMDIDLEANFEKINELKEINTNQEALKR 120
QY 121 QLHQQLHAAVLRQGHFEPQLAAHTDASERTPLIQ-APGGPHQDLRVNFVAGAVEPKA 179
Db 121 NFLELTTELKTLRKTTQQFFDEMADPDLLSESSLLSEPMGRGTPLRLGFGVAGVINRRI 180
QY 180 PALERLLWRACGFLIASFRELEQPLEHPVTGEPATWTMTFLISYNGEQIGQIRKITCF 239
Db 181 PTFERMLRVCKGNVFLRQAEIENFLEDVTDGYVHKSVFIFFQDQLKRVKVKICEGF 240
QY 240 HCHVFPFQEQEARLQALQOQQOQEQELQEVLTGETERFLSVGLRVQLLPQGVQVHKM 299
Db 241 RASLYPCETPQERKEMASGVNTRIDDLQWLNVNQTHEDRQRVLOAAKXIRVFWFKVRM 300
QY 300 KAVYLALNQCSVSTHKCLIAEAWCSVRDLPALQELARDSSMBEG--VSAVAHRIPCRDM 357
```

```
Db 301 KAIYHTLNCNIDVTQKGLIAEWCPCVTDLDSIQFALRGTSSTVDSILNRMTQNT 360
QY 358 PPTLIRNFTASFGQIVDRYGVGRYQEVNPAPYTIITPPFLFAVMFGDVGHGLMLFLPA 417
Db 361 PPTYNKTNKFTYGFQINVDAYGIGTYREINPAPYTIITPPFLFAVMFGDLGHGLMTLPA 420
QY 418 LAMVLAENRPVAKAAQNEIWTFFPRGRVLLLLMGLFSYTYGTGYNECPRSATSIIPSGHS 477
Db 421 VMVLKESRIILSQKNENEMFSTIFSGRYIILLMGVFSYTYGTGLYNDCKSKSLNIFGSSWS 480
QY 478 VAAAMANQSGSDAFILAQHTMLTLDPNVTGVFLGPYPFGIDPTWSLAANHLSPNSPKMKM 537
Db 481 VRPFEDIYNTTEETLRGNPVLQNLNPAVTVGVFGCPYFGIDPTWNTATNKLTLFNSFKMKM 540
QY 538 SVILGVHMAFGVVLGVFNHVFHFGQRHLLLETLPETLFLGLFGYLVFLVYIKMLCVMA 597
Db 541 SVILGIIHMLFGVLSLEFNHTYFKPLNIYFGIPEIIFMTSLFGYLVILIFYKWTAYNA 600
QY 598 ARA-ASPSILHIFINMFLFSPHS-NRLLYPRQEVQATLVVLALAMVPILLIGTPLHL 655
Db 601 KTSEKAPSLHIFINMFLFSGDSQNSMLYSGQKGIQCFLVVVALLCVPMMLLFXPLVL- 659
QY 656 HRHRRRLRRPADROENKAGLLDLPDASVNGSSDEEKAGLLDDEEBAEVLV----- 707
Db 660 --RRQYLRRK-----HLGTLNFGGIRV-----GNGPTDEAEIIOHQDLSTH 699
QY 708 -----PSE-----VLMHQAHTIEFCLGCVSNTASYLRWLWLSLAHAQLSEVLWAM 753
Db 700 SEDAEPTDEVDVFDGDMVHQAHTIEBYCLGICISNTASYLRWLWLSLAHAQLSEVLWAM 759
QY 754 VMRIGLGLRGVGAUVLPDFAAFVMTVAILLVMEGLSAFLHRLHWVEPQNKFSY 813
Db 760 VHIGLGVKSLAG--GLALFFFAAFATLTVAILLMEGLSAFLHRLHWVEPQNKFSY 817
QY 814 GTGYKLSPPF 822
Db 818 GTGFKFLFP 826

RESULT 3
B38656
vacuolar proton pump 116K chain - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 21-Feb-1992 #sequence revision 21-Feb-1992 #text_change 09-Jul-2004
C/Accession: B38656
R.Perin, M.S.; Fried, V.A.; Stone, D.K.; Xie, X.S.; Suedhof, T.C.
J. Biol. Chem. 266, 3877-3881, 1991
A/Title: Structure of the 116-kDa polypeptide of the clathrin-coated vesicle/synaptic v
A/Reference number: B38656; MUID:91139686; PMID:1704894
A/Accession: B38656
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-838 <PER>
A/Cross-references: UNIPROT:P25286; UNIPARC:UPI0000138CED; GB:M58758; NID:g206429; PIDN:
C/Superfamily: vacuolar ATP synthase 95K chain
C/Keywords: alternative splicing; transmembrane protein

Query Match 45.3%; Score 1931; DB 2; Length 838;
Best Local Similarity 46.4%; Pred. No. 1e-132;
Matches 399; Conservative 139; Mismatches 251; Indels 70; Gaps 11;
QY 1 MGSMPRSEVALVOLFLPTAAAYTCVSRIGELGLVEFRDLNASVAFQRRFVVDVVRCEE 60
Db 1 MGELFRSEMTLAQLFLQSEAAAYCCVSELGKGVQFRDLNPDVNVFQKFFVNEVRCEE 60
QY 61 LEKTFPLQEEVRRAGLVLPKGRLPAPPDRDLRIQETERLAQELRDVRGNQOALRA 120
Db 61 MDRKLFVFEKEIRKANIPMDTGENPEVFPFRDMDIDLEANFEKINELKEINTNQEALKR 120
QY 121 QLHQQLHAAVLRQGHFEPQLAAHTDASERTPLIQ-APGGPHQDLRVNFVAGAVEP 176
Db 121 NFLELTTELKTLRKTTQQFFDEMADPDLLSESSLLSEPMGRGAP---LRLGFGVAGVINR 177
```

```
QY 177 HKAPALERLLWRACRGFLIASFRELEOPLHPTGEPATWMTFLISWGSQIOGKIRKIT 236
D 178 ERIPTEFRLMWRVCRGNVFLRQAEIENPLEPDPVTGTVHKSVPFIIFQGDQAKNRVKKIC 237
QY 237 DCFCHVFPLQOEAEALGALQLOQOQOSELQEVLGETERFLSGVLRVQLLPQGOVQV 296
D 238 EGRFASLYPCPETPQERKENASGVNTRIDDLQWLVNQTEDHRQVLOQAANKIRVWFIVK 297
QY 297 HKMKAVYALNQCSTVTHKCLIAEAWCSVRDLPALQEARLDSMERG--VSAAVHRIPC 354
D 298 RKMKAIVHTLNCNIDVTQCLIAEAVCWPTDLDLSIQFALRRGTGHSSTVPSILNRMQT 357
QY 355 RDNPPILIRNRTASFOGIVDRYGVGRYOEVNAPYTIITFPPLFAVMEGDDVGHGLIMF 414
D 358 NQTPPTYNKTKTHGQNTVDAYGIGTYREINPAPYTVITFPPLFAVMEGDDVGHGLIMT 417
QY 415 LPALAMVLAENRPAVKAQNEIWTQFPRGRYLLMLGLFSIYTGFIYNECFSRATSPFS 474
D 418 LFAVWVLRRESRLSOKNENEMFSWPSGRIYIILLMGLFSIYTGFIYNDQCSKSLNIFGS 477
QY 475 GMSVAAMANQSGSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWLSLAANHLISFLNSFK 534
D 478 SWSVRPMFTIGNTEETLLGSSVLQNLPAIPGVFGPGYPIGIDPIWNIATNKLITFLNSFK 537
QY 535 MKMSVILGVVHMAFGVVLGVFNHVFQQRHRLLETLPELTFLGLFGYLVFLVITYKWL 594
D 538 MKMSVILGIIHMLFGVSLFSLFNHYFKKPLNIYFGFIPEIIPFSSLFGYLVILIFYKWT 597
QY 595 VMA-ARAAPSILIHINMFLFSHPS-NRLLYPRQVRVQATLVVLALAMVPIILLGTPL 652
D 598 YDAHSRNPASLIHINMFLFSYPSGNAMLYSGQKIQCFIIVVAMLCVPMWMLLFKPL 657
QY 653 HLLHRRHRLRRPADRQENKAGLLDLPDASVNGSSDEEKAGLDDEEAEALV----- 707
D 658 ILRHQYLRK-----KHLGTLNFGGIRV-----GNGPTEEDAEIIQHDQL 696
QY 708 -----PSE-----VLMHQAIHTIEFCLGCVSNTASYLRWLALAHQASEVL 750
D 697 STHSEDAEPTEDVDFGDMVHQAHTIEYCLGICISNTASYLRWLALAHQASEVL 756
QY 751 WAMVMRIGL-----GLGREGVAAVLPVPIFAAPAVMTVAILVMEGLSAFLHRLH 803
D 757 WTWVHHGLHVRSLAGLG-----LFFIFAATLTVAILLMEGLSAFLHRLH 807
QY 804 WVEFQNKFSYGTGYKLSPPF 822
D 808 WVEFQNKFTGTGFKLFPF 826

RESULT 4
S15795
vacuolar proton pump homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C:Accession: S15795
R:Craxton, M.; Ainscough, R.; Coulson, A.; Dear, S.; Du, Z.; Durbin, R.; Green, P.; Hall
Staden, R.; Sulston, J.; Thierry-Mieg, J.; Thomas, K.; Waterston, R.; Wilson, R.
submitted to the EMBL Data Library, May 1991
A:Reference number: S15786
A:Accession: S15795
A:Molecule type: DNA
A:Residues: 1-935 <CRA>
A:Cross-references: UNIPROT:P30628; UNIPARC:UPI0000175EED; EMBL:Z11115; NID:96953; PID:9
C:Genetics:
A:Interons: 51/3; 110/3; 153/3; 194/2; 264/2; 682/3; 723/3; 770/3; 844/1; 901/2
C:Superfamily: vacuolar ATP synthase 95K chain
C:Keywords: transmembrane protein
```

Query Match 36.1%; Score 1539.5; DB 2; Length 935;
Best Local Similarity 37.0%; Pred. No. 4.3e-104;
Matches 345; Conservative 165; Mismatches 280; Indels 143; Gaps 21;

```
QY 4 MFRSEVALVOLFLPTAAAYTCVSRIGELGVFRDLNANASVAFQRFVVDVWRCEBLEK 63
D 16 IYRSEQMCLAQVLQSDASQCVAELGELGLVFRDLNPDVSSFORQYNEVVRCDMEWR 75
QY 64 TFFTLOEVRACGLVLPKPGRLPAPPRDLLRIQEETERLAOELRDVRCNQOALRAQLH 123
D 76 KLYRLEREIKKOI PMLDTGENPDAPLPREMIODEATFEKLENELEWVKNQNEETLKKQFS 135
QY 124 QLQHLAAVLR-----QGHEPQLA-AAHTDGAERTPLLOAP-----GGPHQD-----LR 166
D 136 ELTELKHILRKTQTFEEHEDMIASSAESGICE--VLSADBEELSGRFS DAMSPKLQ 192
QY 167 VNFVAGAVPHKAPALERLLWRACRGFLIASFRELEOPLHPTGEPATWMTFLISWGE 226
D 193 LRFVAGVIQERLPAFERLLWRACRGVFLRTSEIDVLDNDVTGDPVNRKCVFIIFQGD 252
QY 227 QIOGKIRKKTDCPHCHVFPE---LQOEAEALCALQLOQOQOSELQEVLGETERFLSOVL 282
D 253 HLKTKVKKICEGFRATLYPCPTPQERREMSIGWTRI-----EDLKTVLQOTQDHRHVL 308
QY 283 GRVLQLLPPGQVQHKKAVYALNQCSTVTHKCLIAEAWCSVRDLPALQEARLDSME 342
D 309 VAASKNVRMWLTKVRKIKSIYHTLNLFNIDVTQCLIAEAVWCPIAELDRKMAKRGCTDE 368
QY 343 EG--VSAAVHRIPCRDMPPTLIRTNRTASFOGIVDRYGVGRYOEVNAPYTIITFPPLF 400
D 369 SGSQVPSILNRMETNEAPPYTNKTKFQFQINVDAYGIATYREINPAPYTWISFPPLF 428
QY 401 AVMGDVGHLGLMFLPALAMVLAENRPAVKAQNEIWTQFPRGRYLLMLGLFSIYTGFI 460
D 429 AVMGDMGHGAILMLAALFFILKEQLEAARIKDEIFQTFPGGRYVIFLMGARSITYTGF 488
QY 461 YNECFSRATSIIPSGW-----SVAAMANQSGWSDAFL-----AQHTMLTLDPNVTGVFLG 510
D 489 YNDVFSKINTFGSSWQNTIPESVI-----DYVLDDKEKSESQILPPE-TAFDGN 538
QY 511 PYPFGIDPDTWSLA-ANHLISFLNSFKMMSVILGVVHMAFGVVLGVFNHVFQQRHRLLE 569
D 539 PYPIDGVDPMNLAEKNLSFLNSMKMMSVILFGIAQMTFGVLLSYQNFIFYKSDLDIKYM 598
QY 570 TLPELFLGLGFLGVFLVLYKWLVCVWAARA-----ASPSLIHIFINMFLFSHS 618
D 599 FIPQMIFLSIFLYICILSKWLFPFAGVGTGVLGVKPGSNCAPSLILGLIINFMWKS 658
QY 619 PSN-----RLLYPRQEVVQATLVVLALAMVPIILLGTPLHLLHRRHR--- 661
D 659 NAGFVDDSGETYPQCYLSTWYPCQSFETIFVLVAIACVPMVLFKPYFLWKEERREG 718
QY 662 -----LRRRPADRQENKAGLLDLPDASVNGWSSD 691
D 719 GHRQLATIBIILVVLALVQVPIMLFAKPYFLYRR--DKQOSRY-----STLTAESNQ 768
QY 692 EEKAGGLDDBEAEALV-----PS-----EVLHQAIHTIEFCLGCVSN 729
D 769 HQSVRADINQDDAEVHVAPEQTPKPSGHGHGHDGPLEMGDVMVYQAIHTIEFVLCGVSH 828
QY 730 TASYLRWLALSLAHAQLESEVLWAMVNRIGLGLGREYGAAVLVLPVIFAAPAVMTVAILL 789
D 829 TASYLRWLALSLAHAQLESDVLTMTVFNAPVLDGYTGAIATYI--LFFIFGSLSVFLVL 886
QY 790 MEGLSAFLHALRLHVVFEFQNKFSYGTGYKLSPPF 822
D 887 MEGLSAFLHALRLHVVFEFQNKFSYGTGYKLSPPF 919

RESULT 5
T18565
probable H+-exporting ATPase (EC 3.6.3.6) vacuolar [similarity] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18565; T19177
R:Northmore, B.
submitted to the EMBL Data Library, December 1998
```

A:Reference number: Z18977
A:Accession: T18565
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-865 <WIL>
A:Cross-references: UNIPROT:Q17660; UNIPARC:UPI000007E97E; EMBL:AL031269; PIDN:CAA20334.
A:Experimental source: clone VM02B12L
R:Colles, L.
submitted to the EMBL Data Library, August 1995
A:Reference number: Z19084
A:Accession: T19177
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-865 <W12>
A:Cross-references: UNIPARC:UPI000007E97E; EMBL:Z50872; PIDN:CAA90758.1; GSPDB:GN000020;
A:Experimental source: clone C05D12
C:Genetics:
A:Gene: CESP:VM02B12L.1
A:Map position: 2
A:Introns: 66/1; 181/2; 576/2; 648/3; 775/1
C:Superfamily: vacuolar ATP synthase 95K chain
C:Keywords: hydrolase

Query Match 35.0%; Score 1491.5; DB 2; Length 865;
Best Local Similarity 37.0%; Pred. No. 1.2e-100;
Matches 330; Conservative 165; Mismatches 285; Indels 113; Gaps 18;

QY 1 MGSERSEVALVOLFLPTAAAYTCVSRIGELGLVEFRDLNASVAFQRRFVVDVWRCEE 60
Db 1 MGSYRSEHMKLCQIFFOSEAYQCVAEIGLMAQFIDLNEEQAYTRKFNEVRRCD 60
QY 61 LEKTFTEFQEEVRRAGLVLPPEKGRLPAPPDRDLRIQEBTERLAQELRDVRGNOALRA 120
Db 61 MERKINFEDEITKDLVIPDYDEHPAPQPKHMEANEKLEELVQINKCKVLXN 120
QY 121 QLHQQLHAVALRQGHPEQLAAHT--DGASERTPLAQPGP-----HQDL 165
Db 121 NHVQLLEKMAVLE--HVTSLDDPHSKREAAIMSISAARGEAGPISFGMKDEFDPKVDK 178
QY 166 RVNFVAGAVEPHKAPALERLLWRACRGFLIASFRLEQPLEHPVTGEPATWTFLLISWG 225
Db 179 ELKFTVGVVGRSKAFAFERFLRLSRKVFAPKIQIQOTE--LFSNEFEDKCVFLTFSG 237
QY 226 EQIGKIRKIDTCFH--CHVFPLQOEERLQALQOQSOELQOEVLGETERFLFSQVLG 283
Db 238 EQLRAVKKICDGFQAKCVTPDENPAETKL--LLNIKVTQDMKAVIEKTLDYRSKCIH 295
QY 284 RVQLLPPGQVQVHKKAVYALNOCVSTTHKCLIAEAWCSVRDLPALQEARLDSMBE 343
Db 296 AAATNLKRWGIMLLKLSIFHTLNNFSDVDTOKCLIAECWVPEADIGQVXNSLHMGTHS 355
QY 344 G--VSVAHRIPCRDMPTLIRNRTASFGQIVDYGVRGOEVNPAPYTIITPFLPA 401
Db 356 GSTVPAIUNEMETDKYPTFKLNFQGFQVQFQVQFQVQFQVQFQVQFQVQFQVQFQV 415
QY 402 VMFGDVGHGLMFLPALAVLAENRPVAKAAQNEIWQTFPRGYLLLMGLFSYITGYI 461
Db 416 VMFGDAGHGIIMLIAASAFVFEKLIISMKIKDEIFNTFFGGRVYVLLMGHFAITGYI 475
QY 462 NECFSRATSIIPSGH-----SVAMANQSGWSDAFLAQHTMLTLDPNVT--GVFLGYP 513
Db 476 NDFYSKSNIFGSSWVNPYNQTLANMDAQGADSNITLS---LTFPEIAPNHDYGPYP 531
QY 514 FGIDPIWLAANHLFSFLSKMGMVILGVHMAFGVVLGVNPHVHFGORH--RLLETL 571
Db 532 FGVDPMNLAINRLNFMNKMKTSLIGISQAFGIMLSLNNH1--GNRSVDIVFVFI 589
QY 572 PELTFLGLFGYLVFLVYIKWLCWAARA-----ASPSLIHFINMFL----- 614
Db 590 FOCLFLGCI FVYLCVQLVLMKMIFFVVKPAYIFGRLYPGSNCAPSLIGLINFMVYKSRDA 649
QY 615 -FSHSPSN-----RLLYPRQEVVQATLVVLALAMVPILLG 649

Db 650 SPAHDVGTAAAGKEWVIVNGQNVYTTINDQCYLQWYPNQSLVELILLIIVAVVPVMLLV 709
QY 650 TPLHLLHRRRLRRPPADRRQENKAGLLDLPDASVNGWSSDEEKAGGLDDDEEAEELVPS 709
Db 710 KPFYIRWRHRSGLH-----IDL-----GHGPDHGEFNF 739
QY 710 EYLMHQAHHTTEFCGVCVNTASYLRLWALSALHAQLSEVLWAMVIRIGLGLGREGVAA 769
Db 740 DIMVHQAHHTTEFCVSHVNTASYLRLWALSALHAQLSDVLWMLRMSLTWGGGSGAA 799
QY 770 VVLVPIFAFAVMTVAIILVMELGSAFLHALRHVVEFQNFQKFSYGTGYKLSPF 822
Db 800 ITIL-FYFIFISILSVICILIMELGSAFLHALRHVVEFQNFQKFSYGTGYLQFEFP 851

RESULT 6
T16282
hypothetical protein F35H10.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16282
R:Woessner, J.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F35H10.
A:Reference number: Z18490
A:Accession: T16282
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-873 <W0S>
A:Cross-references: UNIPROT:Q20072; UNIPARC:UPI000008163D; EMBL:U40934; NID:gl072149; P:
C:Genetics:
A:Gene: CESP:F35H10.4
A:Introns: 39/3; 709/3; 785/1; 843/2
C:Superfamily: vacuolar ATP synthase 95K chain

Query Match 33.0%; Score 1406.5; DB 2; Length 873;
Best Local Similarity 36.8%; Pred. No. 1.9e-94;
Matches 329; Conservative 158; Mismatches 303; Indels 103; Gaps 23;

QY 1 MGSERSEVALVOLFLPTAAAYTCVSRIGELGLVEFRDLNASVAFQRRFVVDVWRCEE 60
Db 1 MGSLSRSEMRFCOLIVEKDAAFNIVAEIGKQPVYQVKDLNPNVNSFORTFVKDIRRDE 60
QY 61 LEKTFTEFQEEVRRAGLVLPPEKGRLPAP-----PPRDLRIQEBTERLAQELRDVRGQ 115
Db 61 MERKLRFLSQVVDKDVIP--GRVDTGDTYILTPTSEINTLEGTLEKDKDKVMNDS 117
QY 116 QALRAQLHQLQHAVALR-----QGHEPQLAAAHDTGASERTPLAQPGPHQDLRVNF 169
Db 118 SQLKANFMDLKEDAVLDKTDPEFGGVDDQAELENDDEGAVPRVEKGP-----VNY 172
QY 170 VAGAVEPHKAPALERLLWRACRGFLIASFRLEQPLEHPVTGEPATWTFLLISYGEQIG 229
Db 173 LVGIRIRRELNGFERVLWRACHHTAYIRSSDIEELEDPGTGEKVHKSVFIFLKGDRMR 232
QY 230 QKIRKITDCFHVF-----PFLQOEERLQALQOQSOELQOEVLGET--ERFLSQVL 282
Db 233 SIVEKVCDFKAKFKNCPKTFKERSAR-----NDVRARIQDLQVLQGTREHRT----- 283
QY 283 GRVLQLLPPGQ-----VQVHMKAVYALNOCVSTTHKCLIAEAWCSVRDLPALQEARLD 338
Db 284 -RVLQAAANNHQLKQVRMIKTVFHMLNFTFDGIGRFFVGEWCWIPLKHVEDVRKAIEV 342
QY 339 SSMEEG--VSVAHRIPCRDMPTLIRNRTASFGQIVDYGVRGOEVNPAPYTIITP 396
Db 343 GAERSGSVKPVNLITSVTPYTNKFTAPVQGVQVDSYGIATYATYELNPAPYTIITP 402
QY 397 PFLFAVMDGVGHGLMFLPALAVLAENRPVAKAAQNEIWQTFPRGYLLLMGLFSY 456
Db 403 PFLSCMGDGLGHGCIIMLAGLWFLVRKKNLQARNIKDEIFNMFPGGYIILLMGLFSIH 462
QY 457 TGFYNECFSRATSIIPSGWSVAAMANO--SQWSDAFLAQH---TMLTLDP-----NVTGVF 508

463 AGIINDMFAKSNIFGSGWKNPNYASEIEGWNR--TEHCKEMLVELAPEDAYDHAG-- 518
509 LGPYPPGIDPIWLSAANHLSPNFKMKMSVILGVVHMAFGVILGVFNHVFHFGQRHLL 568
519 -GPYSFGVDPIWIAENKLNFLNKMMLSVILGISQMTFGVILSFNHTYKSKIDIFT 577
569 ETIPELFLGLGLGFLVFLVLYKLCVWAARA-----ASPSILIHFINMPLFSH 617
578 VFIPQMLFMGICFMYLCQILKWLFWTKEATVFGQIYPGSHCAPSLLGLINMFMKD 637
618 SPNSRL--YPRQEVVQATLVVLALAMVPIILLIGTPLHLHR 657
638 RNAGFVVDGKNGEXREVEYCYLSQYVPGSVLEMLIVIAVLCVPMLEGGPIH--HV 695
658 HRRRLRRPADRQENKAGLL-DLPDASVNGWSSDEBKA-----GGDDDEEAEALVPE 710
696 MOOKKAKELHGNATVRANVVSSEIYLVNGGSKKEGAHEEHGGHGEDSEF-----GD 750
711 VLMHQAHTTEPCICGCVSNATSYLRWALSIAHAQLSEVLWAMV-MEIGLGLREVGVAA 769
751 IMVHQAHTTEYVLCVSHSTASYLRWALSIAHAQLSEVLWAMVVFVGGIGISGTAGFIA 810
770 VLVLPVFAAFVMTVAILLVMEGLSAFLHALRLHWHVEFQNKFGYSGTGKLSPF 822
811 VYVV--FFIFVLTISILVMEGLSAFLHALRLHWHVEFQNKFGYSGTGKLSPF 861

RESULT 7
T19492
hypoetical protein C26H9A.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T19492
R:Smyle, R.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z19131
A:Accession: T19492
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1236 <WIL>
A:Cross-references: UNIPROT:Q9XTS8; UNIPARC:UPI000017B7C6; EMBL:Z99169; PIDN:CAB16306.1
A:Experimental source: clone C26H9A
C:Genetics:
A:Gene: CRSP:C26H9A.1
A:Map position: 4
A:Introns: 50/2; 107/2; 119/3; 151/1; 209/2; 327/3; 432/3; 451/3; 482/2; 524/3; 552/2; 5

Query Match 32.2%; Score 1375; DB 2; Length 1236;
Best Local Similarity 35.8%; Pred. No. 6.1e-92;
Matches 334; Conservative 148; Mismatches 321; Indels 130; Gaps 20;

3 SMPSREVALVOLFLPTAAAYTCVSRIGELGLVEFRDLNASVSFAQRRFVVDVVRCEELE 62
291 SMPSRDPMKLYQMILYKEAFCEVABEIGKHGNQVFDLNKMSLYSTFTVKQRCCEME 350
63 KTTFTFLOEEV--RRAGVLVPPKPG----RLPAPPPRDLRLKIQETERLAQELDRVGRNQ 116
351 RKLRFLEKQVITCKPGL---DPKSIDYDLSAPTQAEMLQLEHKLQDLEREFDLNNNDY 407
117 ALRAQLHQLHAALVLR-----QGHPEOLA-----AAHTDG-----ASE 150
408 ALRKNLSNKEFFQVWRLVDEFFQVHKKEBAKARFERSATTDIDMFSSKSGFGGLPSSN 467
151 RTPLLQAPGPHQDLRVNFVAGAVEPHKAPALERLLWRACRGFLIASFRLEQPLEHPVT 210
468 EMFLTPILG---SDDNANFVAGVLPDLKKESFERVLWRACRTAFVRTSDASTVNDPVT 524
211 GEPATWMTPLISVWGBOIGQIKIRKIDCFHCHVFPFLQOBEARLQALQOQSQBELQEV 270
525 LEPLQRCVFTVFFKGSRLRIVEKVCDDGNATQYPCPKSSKDRKMKMSETEGRMDLTVV 584
271 LGETERFLSVLGRVLQLLPPGQVQVHKMKAVYALNQCSTVTHKCLTAENWCSDRLDP 330

585 IDTTQTHRYTILKDMSPFIPIWLKNIQIQKSVFAVMNFTVD-TNGFLAGSECTPAABED 643
331 ALOEALRDSMEEG--VSVAHRIPCRDMPTTIRNRTASTASFOGIVDRVGVGRQYQVNP 388
644 DVRAQLHDGFKASCTEVEPIINELWNAWPPPTFHTNKNFTNVFQSVDSYGVSCVCEVNP 703
389 APTTIITFPFLPFAVFGDVGHLMLFLFALAMVLAENRPAVKAQONRIMWOTFFRGRYLL 448
704 APTTIITFPFLPFAVFGDAAHAILLAALFFIRNERKIESKIRDEIFNTFYGRYIMM 763
449 LMGLFSYTCFTYNECFSRATSIPPSGWSVAAMANQSGWSDAFLA-----OHTMLTLD 502
764 LMGIFSYTCFTYNDAPAKSFNVFGSGWSNSYNETQLDW---WIARYRKHREYSLELVP 820
503 NVTGVFLGPPGIDPIWLSAANHLSPNFKMKMSVILGVVHMAFGVILGVFNHVFHFGQ 562
821 EKSPDIEKTYPPFGVDPIWNIADNRLSFLNSMKMKASVLIIGITQMTFGVFLSVLNHIHFKS 880
563 RHRLLLETLBELFLGLGFLVFLVLYKLCV-----WAARAASPSLIHFIN 611
881 YIDIISNFIPOVIFLSCIFITLCIITVKKWIFFSVNAENVEFPGSHCAPSLLIGLIN 940
612 MFLFSHSPSNRL-----LYPRQEVVQATLVVLALAMVPIILLIGTPL----- 652
941 MFMFKRNEGLNENGEVSYNCHLGYWPNQRLVETILISISLACIPIMLFGKPLWRFV 1000
653 ----HLHRRH--RRLRR--PADRQENKAGLLDLPDASVNGWSSDEKAGGLDDEE 702
1001 TSKRHLKQENKLSKSLRRNGTTVSAPTSPVWDAGPRFEDAE--LLADE-----LDIGE 1053
703 EABLVSPEVLHQAIHTIEPCLCGVSNATSYLRWALSIAHAQLSEVLWAMVVRIGLGLG 762
1054 DIHSLSDIPVHQAIHTIEFVLGVSHSTASYLRWALSIAHAQLSEVWMMHVLITQIHTV 1113
763 REV--GVAAVLVPI-----PAAFAVMTVAILLV 789
1114 DHIENTIAMCLKPPVACVGYFSASAIFFPCLTSLLYGKTYEKEKAPFFIFASUSLSILIM 1173
790 MEGLSAFLHALRLHWHVEFQNKFGYSGTGKLSPF 822
1174 MEGLSAFLHALRLHWHVEFQNKFGYSGTGKLSPF 1206

RESULT 8
H84600
probable vacuolar proton-ATPase subunit [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: H84600
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
eaus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H84600
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-821 <STO>
A:Cross-references: UNIPROT:Q9SJ77; UNIPARC:UPI000000A1530; GB:AE002093; NID:94567273;
C:Genetics:
A:Gene: At2g21410
A:Map position: 2
C:Superfamily: vacuolar ATP synthase 95K chain

Query Match 31.2%; Score 1331; DB 2; Length 821;
Best Local Similarity 37.3%; Pred. No. 5.8e-89;
Matches 317; Conservative 142; Mismatches 306; Indels 84; Gaps 20;

4 MPRSEVALVOLFLPTAAAYTCVSRIGELGLVEFRDLNASVSFAQRRFVVDVVRCEELE 63
17 LMRSEPMQLVQVIVPMESAHLTVSYLGLDLGLVQFKOLNSEKSPFQRTYAAQIKRCGEMAR 76

| | | | |
|----|-----|---|-----|
| Qy | 223 | YWGSQIGOKIRKIYDCHFHCVPPFLQOEAREALGALQLOQSOEQLQEVIGETERFELSQVL | 288 |
| Db | 240 | SHGLDIIKIRIKIAESLDANLIYDVDSNNEGRSSQQLAKYNKQLSDLYTVLTKTTSTTLESEL | 299 |
| Qy | 283 | GRVLQLLPPGQVQVHKMKAVYLAALNOCVSTTHKCLIAEAMCWSVRDLPALQEARLDSMS | 342 |
| Db | 300 | YAIAKELDSWFQDVTREKAIPEILINKSNYDNRKILIAEGHIPDELATLQARLEGMAT | 359 |
| Qy | 343 | EG--VSAVHRIPCRDMPPTLIIRNRPASFGQIVDRYGVGVQYQVNPAPYTIITFPFLF | 400 |
| Db | 360 | LGIDVPSIIQVLDNTHNPTPTPHRTNKPAGFQSDCYGIAQYREINAGLPTIVTFPPMF | 419 |
| Qy | 401 | AVMEGDVGHGLLMLEFALAMVLAENRPAKQAQNEIWQTFPRGRVLIILMLGLFSYITGEI | 460 |
| Db | 420 | AIMFGDMGHGFLMTLAALSVLNEKK-INKNKRGIEIDMAFTGRYIIILMGVFSMYTGFL | 478 |
| Qy | 461 | YNECFSRATSIFFPSGWSVAAMANGSGWSDAFLAQHTMLTLDPNVTGVFLGPYPFGIDPIW | 520 |
| Db | 479 | YNDIFSKTMTIFKSGWK-----WPDHW-----KKGESITATSVGYTPIGLDWM | 522 |
| Qy | 521 | SLAANHSLFNSFKMKSVILGVVHMAFGVVLGVFNHVFQQRHRLLLLETPELPFLGL | 580 |
| Db | 523 | HGTENALLFNSYKQKLSILMGFTHTMTSYEFSLANHLFYNSMIDIIGNFIPGLLFMOGI | 582 |
| Qy | 581 | FGYLAVLVIYKWLCVWAARA-ASPSILIHFNIMELFHSHPNRLLYPPQEVQATLVLA | 639 |
| Db | 583 | FGYLSVCIVIKWADVMDKQKPAFLGNMLNIMEL-SPGTIDDELYPHQAKVQVFLMLMA | 641 |
| Qy | 640 | LAMVPIILLGTPLHLHRRRRRRRRPADROEENKAGILLDLPDASVNGWSSDEE---- | 694 |
| Db | 642 | LVCIPWLLLVKPLHFKFTHKKKSH-----PLPSTEADASSEDELAQQLIS | 687 |
| Qy | 695 | AGGLDDBEEAEALVP-----SEVLHQAIHTIEFCLGCVSNTASYLRLWALS LAHAQLS | 747 |
| Db | 688 | AMDADDAEEBEVSGSGHGDFGDMIHQVHTIEFCLNCVSHSTASYLRLWALS LAHAQLS | 747 |
| Qy | 748 | EVLWAMVMRIGLGLRGVGVAAVVLPIFAFAVMVAAILLVMEGLSAFLHAIRLHWVEF | 807 |
| Db | 748 | SVLWMTIQTAFGRPGVGV--FMVALFAMWFALTCAVLVLMEGTSAMLHSLRLHWVES | 805 |
| Qy | 808 | QNKFPYSGTGKYLSPF 822 | |
| Db | 806 | MSKFFVGEGLPYEPF 820 | |

RESULT 11

T37787

probable vacuolar atpase subunit - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T37787

R:Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V. submitted to the EMBL Data Library, September 1995

A:Reference number: Z21746

A:Accession: T37787

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-805 <Oli>

A:Cross-references: UNIPROT:O13742; UNIPARC:UPI0000138CD1; EMBL:Z98529; PIDN:CF

A:Experimental source: strain 972h-; cosmid c16E8

C:Genetics:

A:Gene: SPDB:SPAC16E8.07c

A:Map position: 1

A:Introns: 14/3

C:Superfamily: vacuolar ATP synthase 95k chain

Query Match 30.08; Score 1280.5; DB 2; Length 805;

Best Local Similarity 36.58; Pred. No. 2.7e-85;

Matches 308; Conservative 122; Mismatches 321; Indels 93; Gaps 18

Qy 26 VSRIGELGLVEFRLDNASVSFAQRRFVVDVVMRCSELEKTFTFLOEVRRAGLVLP----P 81

Db 1 MSALGELSTIHPKDLNPDVYAFQSFVREIRRLDTERLLRYLHSEIDLNGIHPDNLDP 60
QY 82 P--KGRLPAPPRDLRLRIQETETERLAQELRDVGNQALRAQ-LHQQLHAHV----- 131
Db 61 PSYSVLESSTIEDII---ERITREARVRQVLESSQLLEARYLQLOLEFANVLTAKADAPF 117
QY 132 -----LRQHEPQLAAAHDTGASERTPLQ-----APGPHQDLRVNF 169
Db 118 SKSGNTVDPLRNNYETSSIFSGEDDTT--APLIENALELGTGTGTPDSETSFQMTTILDF 175
QY 170 VAGAVEPHKAPALERLLWRACRGFLIASPRELEQPLEHPV-TGEPATWMTFLISYWGEOI 228
Db 176 VSGIITPVKQFLEIRLMTLGNLFIHOVRADDSLHGAENKEBKT--IFLVIAHGTQI 233
QY 229 GOKIRKIDTDPCHVPPPLFQEEARLGAQLOOQOQOELOEVLGETERLSQVLGRVLQ 288
Db 234 LURIRKISLSGATLFFVEEDAPGTSIQOANVSISDLNVALENTSRALYTELTFIASH 293
QY 289 LPFGQVQVHKMAVYLAALNQCSTVTHKCLIAEAWCSVRDLPALQELARDSS--MBEGVS 346
Db 294 ISAMEAVLHKDTVQVMNLFYDQNHKCLIAEGWCPTANLPMVQKTLNISDLTDSQAP 353
QY 347 AVAHRIPCDMPPTLIRNRTASTFGIVDRYGVGRYQVBNPAPYTIITFPFLFAMFGD 406
Db 354 TILNVHTSEQPTVFRVKNKTEGFSIISYGIATYREVNHGIVAIVTFPELFAIMFGD 413
QY 407 VCHGLLMFLFALAMVLAENRPAKAAQNEIMQTFPRGRVLLMLGLFSIYTGFIYNECFS 466
Db 414 LGHGAIMASVALMFLVYKTLGAKKDLDEIVGMVFYGRYIVLLMLGFSYGVFNDDJFS 473
QY 467 RATSIFFPGSWSVAAMANOSGM---SDAFLAQTMLTLDPNVTGVFLGYPFGIDPTWSLA 523
Db 474 KPMISFSSRW-----WPKSEEAJAR-----AVQVGYPIGIDPTWHA 513
QY 524 ANHLSFLNSFKMKSIVLGVMHMGVILGVNHNHFGQHRLLLETPPELFLGLFGY 583
Db 514 DNNLFMSYKMKLSIILGVIHMTCLFLSLNRYFFKRLDIYAVFVPSLIFLEAIFGY 573
QY 584 LVFLVYKWLVCWAAARASPSLIIHFINMFLSHSPSNRLLYPROBVOQATVLAALAMV 643
Db 574 LVITIVKWCIDWKADQPPSLNMLLMLFSPGLEQLPGQKYLQGVILVIAALICV 633
QY 644 PILLGTPLHLHRRRLRRPADRQENKAGLL--DLPDASVNGSSDEKAGGLDDE 701
Db 634 PWLLIVKPVVLRH-----SNEENKYSLSNDLPNV-----DEADALMAVDS 676
QY 702 EBAELP---SEVLHQAHTIEFCLGCVSNSTASYLRMLWALSALAHQSEVLWAMVRIG 758
Db 677 QEKAEPPFELGEVWIHQVHTIEFCLGCVSHTASYLRMLWALSALAHQSSVLWMTLANG 736
QY 759 LGLGREVGVAAVVLPFAAFVMTVAIILLVMEGLSAPLHRLHWHVEFQNKFSYGTGYK 818
Db 737 FMGTGIVG--SIFVILFGFWFIATCVLVAMEGTSAMLHSLRLHWHVEGMSKHFGEGBYA 794
QY 819 LSPF 822
Db 795 FTFF 798

RESULT 12
H84685
probable vacuolar proton-ATPase subunit [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: H84685
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.
Nature 402, 761-769, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H84685
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-780 <STO>
A:Cross-references: UNIPROT:Q9SK06; UNIPARC:UPI00000A9CB0; GB:AB002093; NID:G6598613;
C:Genetics:
A:Gene: At2g28520
A:Map position: 2
C:Superfamily: vacuolar ATP synthase 95K chain

Query Match 28.4%; Score 1210; DB 2; Length 780;

Best Local Similarity 35.3%; Pred. No. 3.6e-80;

Matches 301; Conservative 120; Mismatches 309; Indels 122; Gaps 17;

QY 4 MFRSEVALVOLFLPTAAAYTCVSRIGELGLVFEFDLNASVSAFORFVVDVWRCCELEK 63

Db 13 LMRSEKMTVLQIIIPVESAHSRITVYLGELGLQFRLDLNADKSPFQRTFANQVRCCEMR 72

QY 64 TPTFQEEVRRAGLVLPKGRLPAPPRDLRLRIQETETERLAQELRDVGNQALRAQHL 123

Db 73 KLRFPKQIDKAGLRCSPP---RLETEPDIALGDLERQLADHEHEVLENNSEKLRQTYN 129

QY 124 QLOLHAVALRQ-----GHEPQLAAAHDTGASERTPLQAPG 160

Db 130 ELLEPKIVLEKVGVPAPRLRSISISFLSLIKQLFYSEMNPHGNSQGLR----- 180

QY 161 PHODLRVNFVAGAVEPHKAPALERLLWRACRGFLIASPRELEQPLEHPVTVGPATWMTFL 220

Db 181 -----FISGIINKDKLLKFERMLFATRGNMFLPNQTTSDDEEIMDSTSEWVKVVF 232

QY 221 ISYWGEOIGQKIRKIDTDPCHVPPPLFQEEARLGAQLOOQOQOELOEVLGETERFLSQ 280

Db 233 VFFSGEQARTKILKICEAFGANCYVPEDTTKQRLTREVLRSLSDELTADAGTRHNN 292

QY 281 VLGRVQLQLPQGOVQVHKMAVYLAALNQCSTVTHKCLIAEAWC-----SVRLPALQEL 336

Db 293 ALNSVGYSLTNWITTVRREKAVYDTLNMFLNFDVTKKLVGEGWCPTFAKTQTHEVLRAT 352

QY 337 RDSSMEEGVSVAHRIPCRDMPTLIRNRTASTFGIVDRYGVGRYQVBNPAPYTIITF 396

Db 353 FDSQQGVG--IFHVMQAVESPTTYFRNKLNTNAQEIIDAIGVARYQENNAVSVVY 410

QY 397 PPLFAMFGDVGHGLLMFLFALAMVLAENRPAKAAQNEIMQTFPRGRVLLMLGLFSY 456

Db 411 PPLFAMFGDMGHGICLLIGAL-YLLARERKLSQKLSFMEMLFQGRYVILLMALFSY 469

QY 457 TGFVNECFSRATSIFFPGSWSVAAMANOSGMSDAFLAQTMLTLDPNVTGV--FLGYPYP 514

Db 470 CGLIYNBFFSVFFHIF--GGS--AYKCRDTCSDAY-----TVGLIKYRDPYFP 514

QY 515 GIDPTWSLAAANHLSPNSFKMKSIVLGVMHMGVILGVNHNHFGQHRLLLETPLEL 574

Db 515 GVDPSWRSRTELPYLSNLSKMKMSILLGIAQNNLGLIISFFNARFSGSLDIRYQIFQM 574

QY 575 TFLGLGFLVPLVITYKWLVCWAAARASPSLIIHFINMFLFSHSPSNRLLYPROBVOQAT 634

Db 575 IFLNSLFGYLSLLIIKW-C-----TGSQADLYHIV----- 604

QY 635 LVVLALAMVPIILLGTPLHLHRRRLRRRRPADRQENKAGLLDLPDASVNGSSDEK 694

Db 605 LLLAFIAPVPMWLPFKPFALRKIHMERFQGRTYGVLVSVSEVDLDVEPDSA----- 654

QY 695 AGGLDDEBEAEIVPSEVLHQAHTIEFCLGCVSNSTASYLRMLWALSALAHQSEVLWAMV 754

Db 655 RGGGHHBEENF--SEIFVHQLIHSIEFVLSGVSNTASYLRMLWALSALHSELSTVPEKV 712

QY 755 MRIGLG-----LGREVGVAAVVLPFAAFVMTVAIILLVMEGLSAPLHRLHWHVEFQNK 810

Db 713 LLLAWGYENILTRLIGV-----AVFAFATAFILLMMETLSALFHALRLHWHVEFMGK 763

QY 811 FVSGTGYKLSPP 822

Db 764 FNGDGYKFKPF 775

RESULT 13

T46719
probable vacuolar ATPase (EC 3.6.1.-) proton pump chain 116K [imported] - Leishmania ma
C;Species: Leishmania major
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46719
R;Volckaert, G.; Ivens, A.C.; Lawson, D.; Quail, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, December 1999
A;Reference number: Z23137
A;Accession: T46719
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-775 <VOL>
A;Cross-references: UNIPROT:Q9U150; UNIPARC:UPI000007C98D; EMBL:AL121861; PIDN:CABS9384.
A;Experimental source: strain Friedlin
C;Genetics:
A;Note: L4326.08
C;Superfamily: vacuolar ATP synthase 95K chain
C;Keywords: hydrolase

Query Match 27.9%; Score 1190; DB 2; Length 775;
Best Local Similarity 33.5%; Pred. No. 1e-78;
Matches 285; Conservative 139; Mismatches 301; Indels 126; Gaps 19;

QY 4 MFSESEVALVOLFLPTAAAYTCVSRIGELGLVEFRDLNASVSAPORFVVDVWRCELEK 63
DB 9 LWRSEDMVLSLHMQREVAHDAVLKLGIGQGFEDLNKDVSAFQDFVQEVRRCDMMER 68
QY 64 TFFLOEVRNRRAG---LVLPKPPKGRLPAPPDRLLRIQETRLAQLDRVGNQOALRA 120
DB 69 KLFLOESEKAGVATVDGDAEG-----ETMSSLEHKIDVYSEVELNQYQALIE 121
QY 121 QLHQLQHAVALRQGHPEPQLAAHTDGASERTPLQAPGGPHQDLRVNFVAGAVEPHKAP 180
DB 122 ERNRSEKHEILSR---DFGGATGDC-----VLMVTGVIPKERIP 158
QY 181 ALERLLWRACRGLIASFRELEPHVPTGEPATWMTFLISYWGQIGQKIRKIDTCFH 240
DB 159 LFERLVYRATRGNSIMRTDNIDKPFYNINANEVSVFAYFSAPRLHERLIKIAEANA 218
QY 241 CHVFPLQOEARELQALQLOQSQBQLQELVGLGTERFLSOVLGRVLQQLPPGQVQHKM 300
DB 219 ATVYNTADSEQLTRMHASLQQQVDITQTLNQSAYRQORVLLGIAVCYEWRAVUTEK 278
QY 301 AVYALNQCSVSTHKCLIAEAWCSVRDLPALQEARLDSSMERG--VSAVAHRIPCRDMP 358
DB 279 AVFSTNMMLKFGSG--TAIARGWAPVRSCEIDRTAIAEAYLSGAQVATIEELNTEKTP 336
QY 359 PTLIRNTRFTASFGQIVDRYGVGRYQEVNPAPYTIITFPPLFAVMPGDVGHLLMFLAL 418
DB 337 PSYFKTNKITGSPQSIQVDSYGMARYKEANPGVFVTIITPPLFGVMYGDVGHGILTLFAA 396
QY 419 AMVLA---ENRPVAKAAQNEIWOTFRGRYLLLLMGLFSIYTGFTYNECFSRATSIFFS 474
DB 397 FLVFKESFEGQL---NEIFAMIFGGRYLLLLMGFFAVYMGLLYNDMFGSIFIFAS 451
QY 475 GNSVAAMANQSGWSDAFLAQHTWMLTLDP-NVTGVFLGPYP-----FGIDPIW 520
DB 452 GYR-----WPQ-----LPPEGPDGIVYPSPTGRPSVKPESSVIFGIDSAW 492
QY 521 SLAANHLSFLNSFKMKSIVLGVVMAFGVVLGVFNHVFQQRHRLLETLPLTLGL 580
DB 493 SETENKLEFVNSLKMCSVLIIGVAMQWAGVILSLTNYIYVDSVKVWFVPEVFLSCT 552
QY 581 FGVLVFLVYKMLCVMAARAASPSILIHFMFLFSGHSPSNRL--LVPROEVQVQATLVVL 638
DB 553 FGVMCVLLIIVKMLTTWENTHDAFSLLETWNTFPL---APCTIPLPFGQAALQVMLLV 609
QY 639 ALAWVPILLIGTPLHLHRRHRRRR---PADROEENKAGLLDLFDASVNGWSSDEEK 694
DB 610 SLACVPCMLCVIYVYKKEKHDQKQRAAHPADGEEG----- 648
QY 695 AGGLDDEEAELVPSEVLMHQATHITIEFCLGCVSNTASYLRWLWALSIAHQSEVLWAMV 754

Db 649 -----EDDPQL--SEIIHQIHTIEYVLGCVSNTASYLRWLWALSIAHQSEVFWFA 700
QY 755 MRIGLGLGRGVGAADVLPDFAAFV---MTVAILLVMEGLSAFLHALRLHWVEFQNK 811
Db 701 FLITVDYDSTGIC-----IFFGFAMWMTATIGVLLGMESLSAFLHALRLHWVEFN 754
QY 812 YSGTGYKLSPP 822
Db 755 YAADGYAFEPF 765

RESULT 14

S54554
H+-exporting ATPase (EC 3.6.3.6) chain STV1 - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YMR9796.07; protein YMR054W
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 11-Aug-1995 #text_change 09-Jul-2004
C;Accession: S54554; A54081
R;Devlin, K.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995
A;Reference number: S54548
A;Accession: S54554
A;Molecule type: DNA
A;Residues: 1-890 <DEV>
A;Cross-references: UNIPROT:P37296; UNIPARC:UPI000013615B; EMBL:Z49703; NID:g817880; P
R;Manolson, M.F.; Wu, B.; Proteau, D.; Tailion, B.E.; Roberts, B.T.; Hoyt, M.A.; Jones,
J. Biol. Chem. 269, 14064-14074, 1994
A;Title: STV1 gene encodes functional homologue of 95-kDa yeast vacuolar H(+)-ATPase su
A;Reference number: A54081; MUID:94245725; PMID:7514599
A;Accession: A54081
A;Molecule type: DNA
A;Residues: 1-804,'E',806-890 <MAN>
A;Cross-references: UNIPARC:UPI00001689CE; GB:U06465; NID:g460159; PIDN:AAA20596.1; PID
C;Genetics:
A;Gene: SGD:STV1
A;Cross-references: SGD:S0004658; MIPS:YMR054W
A;Map position: 13R
C;Function:
A;Description: hydrogen ion transport; hydrolase
C;Superfamily: vacuolar ATP synthase 95K chain
C;Keywords: ATP; hydrogen ion transport; hydrolase; transmembrane protein; yeast vacuol
F;451-487/Domain: transmembrane #status predicted <TM1>
F;509-525/Domain: transmembrane #status predicted <TM2>
F;585-600/Domain: transmembrane #status predicted <TM3>
F;614-637/Domain: transmembrane #status predicted <TM4>
F;681-696/Domain: transmembrane #status predicted <TM5>
F;833-851/Domain: transmembrane #status predicted <TM6>

Query Match 26.8%; Score 1144; DB 2; Length 890;
Best Local Similarity 33.2%; Pred. No. 2.8e-75;
Matches 305; Conservative 140; Mismatches 335; Indels 138; Gaps 27;

QY 3 SMRSEVALVOLFLPTAAAYTCVSRIGELGLVEFRDLNASVSAPORFVVDVWRCELE 62
DB 6 AIFRSADMTYQLYIPLVIREVTFLLGKSMFMVMDLNKDLTAFORGVNQRRDEFE 65
QY 63 KTFTFLOEVRNRRAG-----LVLPKPKGRLPAPPDRLLRIQ-----EETERLAQLDR 110
DB 66 RMVGFLEVVVEKHAATWKYILHIDDEGNDIAQPDMDLINTPEPLSLENNVDMVKRITD 125
QY 111 V--RGNQ-----QALRAQLHQL-----QLHAVALRQGHPEPQLAAHTD--- 146
DB 126 CESRARQLDESLSRLSKLNDLLEQRQVIFECSEKFIENVNPGIAGRATNPEIQEERDVDE 185
QY 147 -----GASERTPLQAPGGPHQDLRVN-----FVAGA 173
DB 186 FRMTDPDISETLSDAFSFDDETP--QDRGALGNLDRNOSVEDLSLFLEQYQHYRMITGS 243
QY 174 VEPHKAPELERILWRACRGLIASFRELEPHVPTG-EPATWMTFLISYWGQIGQKIR 232
DB 244 IRRTKVDILNRLWRLRGNLRFQNPPIEPL---LEGKEKVKDCFIIFTHGTELLKKV 300

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2006, 12:59:44 ; Search time 308 Seconds
(without alignments)
2468.710 Million cell updates/sec

Title: US-10-783-519-2
Perfect score: 4266
Sequence: 1 MGSMPFSEVALVQLFLPTA.....HWVEFQNKFGYGTGYKLSPPF 822

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----------------|---------------------|
| 1 | 4249.5 | 99.6 | 830 | 1 VPP3 HUMAN | Q13488 h vacuolar |
| 2 | 4235.5 | 99.3 | 830 | 2 Q8WVC5 HUMAN | Q8WVC5 homo sapien |
| 3 | 3601.5 | 84.4 | 834 | 2 Q91W06 MOUSE | Q91W06 mus musculus |
| 4 | 3600.5 | 84.4 | 834 | 2 Q9JHF5 MOUSE | Q9JHF5 mus musculus |
| 5 | 3597.5 | 84.3 | 834 | 2 Q9JL12 MOUSE | Q9JL12 mus musculus |
| 6 | 3571.5 | 83.7 | 837 | 2 Q216B0 RAT | Q216B0 rattus norv |
| 7 | 2670.5 | 62.6 | 837 | 2 Q918C8 CHICK | Q918C8 gallus gall |
| 8 | 2370.5 | 55.6 | 823 | 2 Q5CZ26_XENTR | Q5CZ26 xenopus tro |
| 9 | 2369.5 | 55.5 | 822 | 2 Q7ZVM7 BRARE | Q7ZVM7 brachydanio |
| 10 | 2295.5 | 53.8 | 827 | 2 Q4R2B2_TETNG | Q4R2B2 tetraodon n |
| 11 | 2188 | 51.3 | 838 | 2 Q4RS23_TETNG | Q4RS23 tetraodon n |
| 12 | 2149 | 50.4 | 854 | 1 VPP2 BOVIN | Q97681 bos taurus |
| 13 | 2145 | 50.3 | 856 | 1 VPP2_MOUSE | P15920 mus musculus |
| 14 | 2145 | 50.3 | 856 | 2 Q216B1 RAT | Q216B1 rattus norv |
| 15 | 2138 | 50.1 | 897 | 2 Q4SJB9_TETNG | Q4SJB9 tetraodon n |
| 16 | 2129 | 49.9 | 490 | 2 Q9CTA9_MOUSE | Q9CTA9 mus musculus |
| 17 | 2119 | 49.7 | 856 | 2 Q6NUM0 HUMAN | Q6NUM0 homo sapien |
| 18 | 2116 | 49.6 | 856 | 1 VPP2 HUMAN | Q9V487 homo sapien |
| 19 | 2068 | 48.5 | 839 | 2 Q918C9_CHICK | Q918C9 gallus gall |
| 20 | 1954 | 45.8 | 838 | 1 VPP1_BOVIN | Q29466 b vacuolar |
| 21 | 1950.5 | 45.7 | 831 | 2 Q8AVM5_XENLA | Q8AVM5 xenopus lae |
| 22 | 1944.5 | 45.6 | 839 | 1 VPP1_MOUSE | Q92194 m vacuolar |
| 23 | 1943 | 45.5 | 832 | 2 Q3TWT5_MOUSE | Q3TWT5 mus musculus |
| 24 | 1941 | 45.5 | 832 | 2 Q216B2 RAT | Q216B2 rattus norv |
| 25 | 1940.5 | 45.5 | 839 | 2 Q216B3 RAT | Q216B3 rattus norv |
| 26 | 1940 | 45.5 | 832 | 2 Q6NXX6_MOUSE | Q6NXX6 mus musculus |
| 27 | 1939 | 45.5 | 838 | 2 Q216B5 RAT | Q216B5 rattus norv |
| 28 | 1938.5 | 45.4 | 845 | 2 Q216B4 RAT | Q216B4 rattus norv |
| 29 | 1935 | 45.4 | 838 | 2 Q918D0_CHICK | Q918D0 gallus gall |
| 30 | 1931 | 45.3 | 838 | 1 VPP1_RAT | P25286 r vacuolar |
| 31 | 1931 | 45.3 | 840 | 1 VPP4_HUMAN | Q9NB54 homo sapien |

| | | | | | | |
|----|--------|------|-----|---|--------------|--------------------|
| 32 | 1930 | 45.2 | 840 | 2 | Q32M47_HUMAN | Q32M47 homo sapien |
| 33 | 1928.5 | 45.2 | 837 | 1 | VPP1_HUMAN | Q93050 h vacuolar |
| 34 | 1927 | 45.2 | 840 | 2 | Q7TIN8_TORMA | Q7TIN8 torpedo mar |
| 35 | 1927 | 45.2 | 846 | 2 | Q6PA83_XENLA | Q6PA83 xenopus lae |
| 36 | 1926.5 | 45.2 | 839 | 2 | Q502H9_BRARE | Q502H9 brachydanio |
| 37 | 1926.5 | 45.2 | 839 | 2 | Q7TIN9_TORMA | Q7TIN9 torpedo mar |
| 38 | 1924.5 | 45.1 | 831 | 2 | Q53ET5_HUMAN | Q53ET5 homo sapien |
| 39 | 1923.5 | 45.1 | 420 | 2 | Q6P735_RAT | Q6P735 rattus norv |
| 40 | 1921.5 | 45.0 | 837 | 2 | Q5RSX1_PONPY | Q5RSX1 pongo pygma |
| 41 | 1920 | 45.0 | 838 | 2 | Q5CZH6_HUMAN | Q5CZH6 homo sapien |
| 42 | 1919.5 | 45.0 | 837 | 1 | VPP1_PONPY | Q5R422 pongo pygma |
| 43 | 1918.5 | 45.0 | 831 | 2 | Q53XT2_HUMAN | Q53XT2 homo sapien |
| 44 | 1907.5 | 44.7 | 835 | 2 | Q4S964_TETNG | Q4S964 tetraodon n |
| 45 | 1907.5 | 44.7 | 837 | 2 | Q5R6N4_PONPY | Q5R6N4 pongo pygma |

ALIGNMENTS

RESULT 1
VPP3 HUMAN
ID VPP3 HUMAN STANDARD; PRT; 830 AA.
AC Q13488: 075877;
DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
DT 01-DEC-2000, sequence version 2.
DT 07-FEB-2006, entry version 49.
DE Vacuolar proton translocating ATPase 116 kDa subunit a isoform 3 (V-ATPase 116-kDa isoform a3) (Osteoclastic proton pump 116 kDa subunit) (OC-116 kDa) (OC116) (T-cell immune regulator 1) (T cell immune response cDNA7 protein) (TIRC7).
GN Name=TCIRG1; Synonyms=ATP6N1C, ATP6VOA3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM LONG).
RC TISSUE=Osteoclastoma;
RX MEDLINE=96158968; PubMed=8579597; DOI=10.1006/bbrc.1996.0145;
RA Li Y.P., Chen W., Staehenko P.;
RT "Molecular cloning and characterization of a putative novel human osteoclast-specific 116-kDa vacuolar proton pump subunit.";
RN [2]
RA Biochem. Biophys. Res. Commun. 218:813-821(1996).
NUCLEOTIDE SEQUENCE (ISOFORM SHORT).
RA Utaku N., Heinemann T., Bulwin C.-G., Beinke S., Beato F., Randall J., Busconi L., Delphire E., Robertson E.R., Kojima R., Volk H.D., Milford B.L., Gullans S.R.;
RA Milford B.L., Gullans S.R.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
NUCLEOTIDE SEQUENCE (ISOFORM SHORT).
RX MEDLINE=99263502; PubMed=10329006; DOI=10.1006/geno.1999.5751;
RA Heinemann T., Bulwin G.C., Randall J., Schnieders B., Sandhoff K., Volk H.D., Milford E., Gullans S.R., Utaku N.;
RA "Genomic organization of the gene coding for TIRC7, a novel membrane protein essential for T cell activation.";
RN [4]
Genomics 57:398-406(1999).
VARIANTS OPTB1 ARG-405 AND LEU-444.
RX PubMed=11532986; DOI=10.1093/hmg/10.17.1767;
RA Sobacchi C., Frattini A., Orchard P., Porras O., Texcan I., Andolina M., Babul-Hirji R., Baric I., Canham N., Chitayat D., Dupuis-Girod S., Ellis I., Etzioni A., Fasth A., Fisher A., Gerritsen B., Gulino V., Horwitz E., Klamroth V., Lanino E., Miolo M., Musio A., Matthijs G., Nonomaya S., Notarangelo L.D., Ochs H.D., Superti Furga A., Valiaho J., van Hove J.B.K., Vihinen M., Vujic D., Vezzoni P., Villa A.;
RA "The mutational spectrum of human malignant autosomal recessive osteopetrosis.";
RN Hum. Mol. Genet. 10:1767-1773(2001).
RL [5]
RP VARIANT OPTB1 ARG-405.

RX PubMed=12552563; DOI=10.1002/humu.10165;
 RA Scineca J.-C., Quincey D., Parrinello H., Romatet D., Grosgeorge J.,
 RA Gaudray P., Philip N., Fischer A., Carle G.F.,
 RT "Novel mutations in the TCIRG1 gene encoding the a3 subunit of the
 RT vacuolar proton pump in patients affected by infantile malignant
 RT osteopetrosis.";
 RL Hum. Mutat. 21:151-157(2003).
 RN [6]
 RP VARIANTS OPTB1 PRO-141; ARG-405; ASN-462 DEL; ASN-517 AND ARG-775.
 RX PubMed=15300850; DOI=10.1002/humu.20076;
 RA Subani L., Pangrazio A., Sobacchi C., Taranta A., Mortier G.,
 RA Savarirayan R., Villa A., Orchard P., Vezzoni P., Albertini A.,
 RA Frattini A., Pagani F.,
 RT "TCIRG1-dependent recessive osteopetrosis: mutation analysis,
 RT functional identification of the splicing defects, and in vitro rescue
 RT by U1 snRNA.";
 RL Hum. Mutat. 24:225-235(2004).
 CC -|- FUNCTION: Part of the proton channel of V-ATPases (By similarity).
 CC Seems to be directly involved in T cell activation.
 CC -|- SUBUNIT: The V-ATPase is an heteromultimeric enzyme composed of at
 CC least thirteen different subunits. It has a membrane peripheral V1
 CC sector for ATP hydrolysis and an integral V0 for proton
 CC translocation. The V1 sector comprises subunits A-H, whereas V0
 CC includes subunits a, d, c, c', and c''.
 CC -|- SUBCELLULAR LOCATION: Membrane; multi-pass membrane protein (By
 CC similarity).
 CC -|- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=Q13488-1; Sequences=Displayed;
 CC Name=Short;
 CC IsoId=Q13488-2; Sequences=VSP 000345;
 CC Note=No experimental confirmation available;
 CC -|- TISSUE SPECIFICITY: The long isoform is highly expressed in
 CC osteoclasts. The short isoform is highly expressed in thymus.
 CC -|- DISEASE: Defects in TCIRG1 are a cause of autosomal recessive
 CC osteopetrosis (OPTB1) [MIM:259700]; also called autosomal
 CC recessive Albers-Schönberg disease. Osteopetrosis is a rare
 CC genetic disease characterized by abnormally dense bone, due to
 CC defective resorption of immature bone. The defect is an
 CC heterogeneous disorder of bone metabolism, which, if untreated, has
 CC a fatal outcome. It occurs in two forms: a severe autosomal
 CC recessive form, occurring in utero, infancy, or childhood, and a
 CC benign autosomal dominant form, occurring in adolescence or
 CC adulthood. However, mild and atypical forms have also been
 CC reported. It is not known whether these latter cases represent
 CC separate nosological entities or a mild form of the recessive
 CC conditions. However, given the large number of genes and gene
 CC mutations associated with this condition a large spectrum of
 CC clinical manifestation is expected as seen in animal studies. The
 CC features of OPTB1 are macrocephaly, progressive deafness and
 CC blindness, hepatosplenomegaly, and severe anemia beginning in
 CC early infancy or in fetal life. Deafness and blindness are
 CC generally thought to represent effects of pressure on nerves.
 CC -|- SIMILARITY: Belongs to the V-ATPase 116 kDa subunit family.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 CC EMBL; U45285; AAA97878.1; -; mRNA.
 CC EMBL; AF025374; AAC35742.1; -; mRNA.
 CC EMBL; AF033033; AAD31081.2; -; Genomic_DNA.
 CC Ensembl; ENSG00000110719; Homo sapiens.
 CC H-InvDB; HTX0009874; -;
 CC HGNC; HGNC:11647; TCIRG1.
 CC MIM; 259700; phenotype.
 CC MIM; 604592; Gene.
 CC GO; GO:005887; C:integral to plasma membrane; TAS.
 CC GO; GO:0005886; C:plasma membrane; TAS.
 CC GO; GO:0005215; F:transporter activity; TAS.
 CC GO; GO:0006968; P:cellular defense response; TAS.
 CC GO; GO:0008284; P:positive regulation of cell proliferation; TAS.
 CC GO; GO:0015992; P:proton transport; TAS.

DR InterPro: IPR002490; V_ATPase_sub116.
 DR PANTHER; PTHR11629; V_ATPase_sub116; 1.
 DR Pfam; PF01496; V_ATPase_1; 1.
 KW Alternative splicing; Disease mutation; Glycoprotein;
 KW Hydrogen ion transport; Ion transport; Membrane; Transmembrane;
 KW Transport.
 FT CHAIN 1 830 Vacuolar proton translocating ATPase 116
 FT kDa subunit a isoform 3.
 FT /FTID=PRO 0000119218.
 FT Extracellular (Potential).
 FT TOPO DOM 1 397 Potential.
 FT TRANSMEM 398 418 Potential.
 FT TRANSMEM 445 465 Potential.
 FT TRANSMEM 504 524 Potential.
 FT TRANSMEM 538 558 Potential.
 FT TRANSMEM 576 596 Potential.
 FT TRANSMEM 636 656 Potential.
 FT TRANSMEM 771 791 Potential.
 FT CARBOHYD 41 41 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 483 483 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 503 503 N-linked (GlcNAc...) (Potential).
 FT VARSPLIC 1 216 Missing (in isoform Short).
 FT /FTID=VSP 000345.
 FT A -> P (in OPTB1).
 FT VARIANT 141 141 /FTID=VAR 020988.
 FT VARIANT 405 405 G -> R (in OPTB1).
 FT VARIANT 444 444 /FTID=VAR 019589.
 FT VARIANT 462 462 R -> L (in OPTB1).
 FT VARIANT 517 517 /FTID=VAR 019570.
 FT VARIANT 775 775 Missing (in OPTB1).
 FT CONFLICT 377 377 /FTID=VAR 020989.
 FT CONFLICT 603 603 D -> N (in OPTB1).
 FT SEQUENCE 830 AA; 92998 MW; 62EAC8A9A22DC698B CRC64;
 Query Match 99.6%; Score 4249.5; DB 1; Length 830;
 Best Local Similarity 99.8%; Pred. No. 1.3e-279;
 Matches 821; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1 MGSMRSEVALVQLPPTAAAYTCVSRIGELGVEFRDLNVSFAFORPVVDVWRCEE 60
 DB 1 MGSMRSEVALVQLPPTAAAYTCVSRIGELGVEFRDLNVSFAFORPVVDVWRCEE 60
 QY 61 LEKTTFTLQEEVRRAGLVLPPEKGRLPAPPPRDLRIQEETERLAQELRDVRGNQALRA 120
 DB 61 LEKTTFTLQEEVRRAGLVLPPEKGRLPAPPPRDLRIQEETERLAQELRDVRGNQALRA 120
 QY 121 QLHQLHAAVLRLQGHFQOLAAAHDTGASERTPLLOAGPGPHQDLRVNPFVAGAVEPHKAP 180
 DB 121 QLHQLHAAVLRLQGHFQOLAAAHDTGASERTPLLOAGPGPHQDLRVNPFVAGAVEPHKAP 180
 QY 181 ALERLLWRACRGFLIASPRELEQPLEHPVTGSPATWMTFLISYWGEOIGQKIRKTKDCPH 240
 DB 181 ALERLLWRACRGFLIASPRELEQPLEHPVTGSPATWMTFLISYWGEOIGQKIRKTKDCPH 240
 QY 241 CHVFPFLQEEARLQALQQLQQSQELQVGLGTERFLSQVLGRVLQLLPQGVQVHKMK 300
 DB 241 CHVFPFLQEEARLQALQQLQQSQELQVGLGTERFLSQVLGRVLQLLPQGVQVHKMK 300
 QY 301 AVYLAHQCSVSTTHKCLIAEAWCSVRDLPALQELALRSSMEEGSVSAVAHRIPCRDMPT 360
 DB 301 AVYLAHQCSVSTTHKCLIAEAWCSVRDLPALQELALRSSMEEGSVSAVAHRIPCRDMPT 360
 QY 361 LIRTNRTASFOGIVDRYGVGRYQEVNPAPYTIITFPFLFAVMFGDVGHGLMFLPALAM 420
 DB 361 LIRTNRTASFOGIVDRYGVGRYQEVNPAPYTIITFPFLFAVMFGDVGHGLMFLPALAM 420
 QY 421 VLAENRPVAKAAQNEIWQTFPRGRYLLLMGLFSYTGFIYNECFSRATSIFFPSGWSVAA 480
 DB 421 VLAENRPVAKAAQNEIWQTFPRGRYLLLMGLFSYTGFIYNECFSRATSIFFPSGWSVAA 480

Qy 481 MANQSGWSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWLSLAANHLSFLNSFKMKMSVI 540
 Db 481 MANQSGWSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWLSLAANHLSFLNSFKMKMSVI 540
 Qy 541 LGVVMHAFVGVVGVFNHVFQHRLLLETLPELTFLGLFGYLVFLVIYKWLVCWMAARA 600
 Db 541 LGVVMHAFVGVVGVFNHVFQHRLLLETLPELTFLGLFGYLVFLVIYKWLVCWMAARA 600
 Qy 601 AS-PSILIHFINMFLFSGHSPSNRLLYPRQEVVQATLVVLALAMVPIILLGTPLHLHRRH 659
 Db 601 ASAPSILIHFINMFLFSGHSPSNRLLYPRQEVVQATLVVLALAMVPIILLGTPLHLHRRH 660
 Qy 660 RLRRRRPADRQENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLHQAIHT 719
 Db 661 RLRRRRPADRQENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLHQAIHT 720
 Qy 720 IEFCLGCVSNTASYLRRLWALSALHAQSEVLWAMVMRIGLGLREVGVAAVVLVPIPAAF 779
 Db 721 IEFCLGCVSNTASYLRRLWALSALHAQSEVLWAMVMRIGLGLREVGVAAVVLVPIPAAF 780
 Qy 780 AVMTVAILLVMEGLSALFLHALRLHWHVFNKFGYSGTGKLSPP 822
 Db 781 AVMTVAILLVMEGLSALFLHALRLHWHVFNKFGYSGTGKLSPP 823

RESULT 2

QBWVC5 HUMAN
 ID QBWVC5 HUMAN PRELIMINARY; PRT; 830 AA.
 AC QBWVC5;
 DT 01-MAR-2002, integrated into UniProtKB/TrEMBL.
 DT 21-FEB-2006, entry version 10.
 DE T-cell, immune regulator 1, isoform a.
 GN Names=TCIRG1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RC TISSUE=lung and pancreas;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RC NUCLEOTIDE SEQUENCE.
 RP TISSUE=pancreas;
 RC NUCLEOTIDE SEQUENCE.
 RG NIH MGC Project;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RC NUCLEOTIDE SEQUENCE.
 RP TISSUE=lung;
 RG NIH MGC Project;

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License

CC
 DR EMBL; BC018133; AAH18133.1; -; mRNA.
 DR EMBL; BC032465; AAH32465.1; -; mRNA.
 DR Ensembl; ENSG00000110719; Homo sapiens.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0015992; P:proton transport; IEA.
 DR InterPro; IPR002490; V ATPase sub116.
 DR PANTHER; PTHR11629; V ATPase sub116; 1.
 DR Pfam; PF01496; V ATPase I; 1.
 DR SQ SEQUENCE 830 AA; 92968 MW; 50751B41B171D9D2 CRC64;

Query Match 99.3%; Score 4235.5; DB 2; Length 830;
 Best Local Similarity 99.6%; Pred. No. 1.2e-278;
 Matches 820; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 1 MGSFMRSEEVALLVOLFLPTAAAYTCVSRGLGELGVEFRDLNASVSATQRRFVVDVRCCE 60
 Db 1 MGSFMRSEEVALLVOLFLPTAAAYTCVSRGLGELGVEFRDLNASVSATQRRFVVDVRCCE 60
 Qy 61 LEKTFELQEEVRAGLVLPKGRLPAPPPRDLRLIQEETERLAQELRDVGRNQOALRA 120
 Db 61 LEKTFELQEEVRAGLVLPKGRLPAPPPRDLRLIQEETERLAQELRDVGRNQOALRA 120
 Qy 121 QLHLQLHAAVLRQGHPEPQLAAAHDTGASERTPLQAPGGPHQDLRVNFVAGAVEPHKAP 180
 Db 121 QLHLQLHAAVLRQGHPEPQLAAAHDTGASERTPLQAPGGPHQDLRVNFVAGAVEPHKAP 180
 Qy 181 ALERLLWRACRGFLIASFRELEQPLEHPVTGEBATWMTFLISYWGEOIGQKIRKITDCFH 240
 Db 181 ALERLLWRACRGFLIASFRELEQPLEHPVTGEBATWMTFLISYWGEOIGQKIRKITDCFH 240
 Qy 241 CHVFPFLQEEEARLQALQQLQOQSOELQEVLGETERFLSVGLRVLQLLPPGQVQVHKMK 300
 Db 241 CHVFPFLQEEEARLQALQQLQOQSOELQEVLGETERFLSVGLRVLQLLPPGQVQVHKMK 300
 Qy 301 AVYLAALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLDSMEEGVSAVAHRIPCRDMPPT 360
 Db 301 AVYLAALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLDSMEEGVSAVAHRIPCRDMPPT 360
 Qy 361 LIRNRTASFGQIVDRYGVGRYQEVNPAPYITITFPFLFAVMPGDVGHGLMLFLALAM 420
 Db 361 LIRNRTASFGQIVDAYGVGRYQEVNPAPYITITFPFLFAVMPGDVGHGLMLFLALAM 420
 Qy 421 VLAENRPAKAAQNEIWTQFRCGRYLLMLGLPSIYTGFIYNECFSRATSIFFPSGWSVAA 480
 Db 421 VLAENRPAKAAQNEIWTQFRCGRYLLMLGLPSIYTGFIYNECFSRATSIFFPSGWSVAA 480
 Qy 481 MANQSGWSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWLSLAANHLSFLNSFKMKMSVI 540
 Db 481 MANQSGWSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWLSLAANHLSFLNSFKMKMSVI 540
 Qy 541 LGVVMHAFVGVVGVFNHVFQHRLLLETLPELTFLGLFGYLVFLVIYKWLVCWMAARA 600
 Db 541 LGVVMHAFVGVVGVFNHVFQHRLLLETLPELTFLGLFGYLVFLVIYKWLVCWMAARA 600
 Qy 601 AS-PSILIHFINMFLFSGHSPSNRLLYPRQEVQATLVVLALAMVPIILLGTPLHLHRRH 659
 Db 601 ASAPSILIHFINMFLFSGHSPSNRLLYPRQEVQATLVVLALAMVPIILLGTPLHLHRRH 660
 Qy 660 RLRRRRPADRQENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLHQAIHT 719
 Db 661 RLRRRRPADRQENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLHQAIHT 720
 Qy 720 IEFCLGCVSNTASYLRRLWALSALHAQSEVLWAMVMRIGLGLREVGVAAVVLVPIPAAF 779
 Db 721 IEFCLGCVSNTASYLRRLWALSALHAQSEVLWAMVMRIGLGLREVGVAAVVLVPIPAAF 780
 Qy 780 AVMTVAILLVMEGLSALFLHALRLHWHVFNKFGYSGTGKLSPP 822

RP NUCLEOTIDE SEQUENCE
RX MEDLINE=20187595; PubMed=10722719; DOI=10.1074/jbc.275.12.8760;
RA Toyomura T., Oka T., Yamaguchi C., Wada Y., Futai M.;
RT "Three subunit a isoforms of mouse vacuolar H⁺-ATPase. Preferential
expression of the a3 isoform during osteoclast differentiation."; J.
RL J. Biol. Chem. 275:8760-8765(2000).
RN [2].
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=129/SV;
RX MEDLINE=20173225; PubMed=10709991; DOI=10.1016/S8756-3282(99)00278-1;
RA Scimeca J.-C., Franchi A., Trojani C., Parrinello H., Grosgeorge J.,
RA Robert C., Jaillon O., Poirier C., Gaudray P., Carle G.F.;
RT "The gene encoding the mouse homologue of the human osteoclast-
specific 116-kDa V-ATPase subunit bears a deletion in osteosclerotic
RT (oc/oc) mutants."; Bone 26:207-213(2000).
RL Bone 26:207-213(2000).
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR ENBL; AB022322; BAA93006.1; -; mRNA.
DR ENBL; AF188702; AAF37193.1; -; Genomic DNA.
DR Ensembl; ENSMUSG0000001750; Mus musculus.
DR MGI; MGI:1350931; Tcirlg1.
DR GO; GO:0016021; C:integral to membrane; RCA.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0008553; F:hydrogen-exporting ATPase activity, phospho. .; RCA.
DR GO; GO:0006818; P:hydrogen transport; RCA.
DR InterPro; IPR002490; V_ATPase sub116.
DR PANTHER; PTHR11629; V_ATPase sub116; 1.
DR Pfam; PF01496; V_ATPase I; 1.
SQ SEQUENCE 834 AA; 93460 MW; E95ECD70C26367C0 CRC64;
Query Match 84.4%; Score 3600.5; DB 2; Length 834;
Best Local Similarity 84.1%; Pred. No. 1.6e-235;
Matches 696; Conservative 51; Mismatches 74; Indels 7; Gaps 5;
F
QY 1 MGSMFSESEVALVQLFLPTAAATCTVSRIGELGLVEFRDLNASVSAPQRRFVVDVNRCEE 60
DB 1 MGSMFSESEVALVQLLPTGSAVNCVSQLGELGLVEFRDLNESVSAPQRRFVVDVNRCEE 60
QY 61 LEKTFTFLOEVRAGLVLPKGRLLPAPPDRLLRIQETRLAQELRDVRGNQQAALRA 120
DB 61 LEKTFTFLEVRAGLTLPAPPGTLPAPPDRLLRIQETRLAQELRDVRGNQQAALRA 120
QY 121 QLHQLQHAVALRQGHPEPQLAAHTDGA-SERTPLLOAPGGPHQDLRVNFVAGAVEPHKA 179
DB 121 QLHQLRLHSAVLGQSHSPVVAADHTGPRSETTPLLPGTRGPHSLDKVNFVAGAVEPYKA 180
QY 180 PALERLLWRACRGFLTASFRELEQPLEHPVTGEPATWMTPLISYGEIQGKIRKITDCF 239
DB 181 ALLERLLWRACRGFLTASFRETEGQLEDPTVTEGEPATWMTFVSYGEIQGKIRKITDCF 240
QY 240 HCVFVFLQOEAEARLQALQOQSQELQEVLCETERFSLQVLRVLLPQGVQVHKM 299
DB 241 HCVFVFLQOEAEARFRLQALQOQSQELQEVLCETERFSLQVLRVLLPQGVQVHKM 300
QY 300 KAVYLTALNQCSTVTHKCLIAEAWCSVRDLPALQALRDSMEEGVSVAHRIPCRDMP 359
DB 301 KAVYLTALNQCSTVTHKCLIAEAWCAARDLPTVQALQSGSSEGVSAVAHRIPCQDMP 360
QY 360 TLIRTRFTASFGIVDRVGVGRYQEVNPAFYTIITPPFLFAVFGDVGHGLLMFLPALA 419
DB 361 TLIRTRFTSSFGIVDAVGVGRYQEVNPAFYTIITPPFLFAVFGDVGHGLLMFLPALA 420
QY 420 MVLAENRPVAKAONEIWQTFGRVYLLLMGLFSYVTGTYNECSRATSIIPSGWSVA 479
DB 421 MVLTENRPVAKAONEIWQTFGRVYLLLMGLFSYVTGTYNECSRATSIIPSGWSVA 480
QY 480 AMANQSGWSDAFIAQHTMLTLDPNVTGVFLGYPFGFDPLWLSLAANHLSPINSFKMKMSV 539
DB 481 AMANQSGWSDXYLSQHSMLTLDNITGVFLGYPFGFDPLWLSLAANHLSPINSFKMKMSV 540

QY 540 ILGVVMAFGVVLGVFNHVFHQHRLLETLBELTFLGLFGVLPVLIYKMLCVWAAR 599
DB 541 ILGVVMAFGVFLSIENHVFHQHRLLETLBELTFLGLFGVLPVLIYKWNVSAAS 600
QY 600 AAS-PSLIHFNNFLFSPNRLLYPROEVQATLVIALAMVPILLAGTPLHLA--H 656
DB 601 ASSAPSLIHFNNFLFSPNRLLYFHGOEVQVYLVVLALATVPLLGLTPLYLRQH 660
QY 657 RHRRLRRPADRQENKAGLLDLPASV--NGWSSDEEKAGGLDDEEAEELVPSEVLMH 714
DB 661 RHRNTQRRPAGQDEDTDKLASPDASTLNSWSDEEKAGSPGD-EETEFVSEIFMH 719
QY 715 QAIHTTFEFCGCVNTASYLRWLWALSLAQSLSEVLWAMVRIGLGLRGREVGAUVLVP 774
DB 720 QAIHTTFEFCGCVNTASYLRWLWALSLAQSLSEVLWAMVRIGLGLRGREVGAUVLVP 779
QY 775 IFAAFVMTVAIILLVMEGLSAFLHALRLHWVEFQNFYSCTGYKLSPP 822
DB 780 VFAAFVMTVAIILLVMEGLSAFLHALRLHWVEFQNFYSCTGYKLSPP 827
RESULT 5
Q9JUL12 MOUSE
ID Q9JUL12_MOUSE PRELIMINARY; PRT; 834 AA.
AC Q9JUL12;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Vacuolar proton-translocating ATPase 100 kDa subunit isoform a3.
GN Name=Tcirlg1; (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=20167151; PubMed=10702241; DOI=10.1074/jbc.275.10.6824;
RA Nishi T., Forgac M.;
RT "Molecular cloning and expression of three isoforms of the 100-kDa a
subunit of the mouse vacuolar proton-translocating ATPase."; J.
RL J. Biol. Chem. 275:6824-6830(2000).
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF218253; AAF59922.1; -; mRNA.
DR Ensembl; ENSMUSG00000001750; Mus musculus.
DR MGI; MGI:1350931; Tcirlg1.
DR GO; GO:0016021; C:integral to membrane; RCA.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0008553; F:hydrogen-exporting ATPase activity, phospho. .; RCA.
DR GO; GO:0006818; P:hydrogen transport; RCA.
DR InterPro; IPR002490; V_ATPase sub116.
DR PANTHER; PTHR11629; V_ATPase sub116; 1.
DR Pfam; PF01496; V_ATPase I; 1.
SQ SEQUENCE 834 AA; 93506 MW; EB8E42B45163004F CRC64;
Query Match 84.3%; Score 3597.5; DB 2; Length 834;
Best Local Similarity 84.1%; Pred. No. 2.5e-235;
Matches 696; Conservative 50; Mismatches 75; Indels 7; Gaps 5;
QY 1 MGSMFSESEVALVQLFLPTAAATCTVSRIGELGLVEFRDLNASVSAPQRRFVVDVNRCEE 60
DB 1 MGSMFSESEVALVQLLPTGSAVNCVSQLGELGLVEFRDLNESVSAPQRRFVVDVNRCEE 60
QY 61 LEKTFTFLOEVRAGLVLPKGRLLPAPPDRLLRIQETRLAQELRDVRGNQQAALRA 120
DB 61 LEKTFTFLEVRAGLTLPAPPGTLPAPPDRLLRIQETRLAQELRDVRGNQQAALRA 120
QY 121 QLHQLQHAVALRQGHPEPQLAAHTDGA-SERTPLLOAPGGPHQDLRVNFVAGAVEPHKA 179
DB 121 QLHQLRLHSAVLGQSHSPVVAADHTGPRSETTPLLPGTRGPHSLDKVNFVAGAVEPYKA 180

```
Db 121 QLHQLRLHSAVLGQSHSPVAAADHTEGPPSETTLLPGTRGPHSDLKUNFVAGAVEPYKA 180
Qy 180 PALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEGIQGKIRKITDCF 239
Db 181 AALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEGIQGKIRKITDCF 240
Qy 240 HCHVPPFLQOEERARLGALQLOQOQOEVLGETERFLS QVLGRVLQLLPPGQVQVHKM 299
Db 241 HCHVPPFLQOEERARLTLLQLOQOQOEVLGETERFLS QVLGRVQQLLPPQVQVHKM 300
Qy 300 KAVYLALNQCSTVTHKCLIAEAWCSVRDLPALQALRDSSEEGVSAVAHRIPCRDMP 359
Db 301 KAVYLTLNQCSTVTHKCLIAEAWCAARDLPTVQALQSGSSEEGVSAVAHRIPCQDMP 360
Qy 360 TLIRNRTTASFOGIVDRYGVRYQEVNPAPTIITFPFLFAMFGDVGHGLMLFALA 419
Db 361 TLIRNRTTSSFOGIVDAVGVRYREVNPAPYTIITFPFLFAMFGDVGHGLMLFALA 420
Qy 420 MYLAENRPAVKAQNEIWTFFRGYLLMLGLFSIYTGFIYNECFSRATSIFPQSGWSVA 479
Db 421 MYLTENRPAVKAQNEIWTFFRGYLLMLGLFSIYTGFIYNECFSRATTIFPQSGWSVA 480
Qy 480 AMANQSGSDAFLAQHTMLTLDPNVTGVLPGYPFGIDPINSLAANHLSFLNSFKMKMSV 539
Db 481 AMANQSGSDEYLSQHSMLTLNPNITGVFLGYPFGIDPINSLATNHLNLSFLNSFKMKMSV 540
Qy 540 ILGVVHMAFGVVLGVFNHVFQORHLLLETLPELTFLGLFGYLVFLVYKWLVCWMAAR 599
Db 541 ILGVTHMAFGVPLSIFNHHVFQORHLLLETLPELTFLGLFGYLVFLVYKWLNSAAS 600
Qy 600 AAS-PSILIHFINMFLFSQNSNRLLYPROEVQVQATLVVLALAMVPIILLGTPHLL--H 656
Db 601 ASSAPSIILHFINMFLFSQNPINLLPHQOEVOVQVVLVVLALATVPILLGTPHLLRQH 660
Qy 657 RHRRLRRRPAQOBENKAGLLDLPDASV--NGWSSDEKAGGLDDEBEAEVLVPSVLMH 714
Db 661 RHRNRTORRPAQOBEDTKLLASPDASTLNSWSPDEEKAGSGD-BETEFPVSEIPMH 719
Qy 715 QAIHTIEFCGCVSNSTASYLRWALSALAHQSEVLWAMVMRIGLGLGREGVAAVVLVP 774
Db 720 QAIHTIEFCGCVSNSTASYLRWALSALAHQSEVLWAMVMRIGLGMGREIGVAAVVLVP 779
Qy 775 IFAPAVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYSGTGKLSPP 822
Db 780 VFAAPAVLTVAILLVMEGLSAFLHALRLHWVEFQNKFSYSGTGKLSPP 827
```

RESULT 6

```
Q216B0 RAT PRELIMINARY; PRT; 834 AA.
AC Q216B0
DT 07-MAR-2006, integrated into UniProtKB/TrEMBL.
DT 07-MAR-2006, sequence version 1.
DT 07-MAR-2006, entry version 1.
DE V-H-ATPase subunit a3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Wistar;
RA Guyon S., Amar M., Fossier P., Morel N.;
RL "Neurons coexpress three different v-ATPase subunit a isoforms.";
RT Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; DQ286426; AB01445.1; --; mRNA.
SQ SEQUENCE 834 AA; 93168 MW; 2E0DCFF9BED3979B CRC64;
```

Query Match 83.7%; Score 3571.5; DB 2; Length 834;

Best Local Similarity 83.7%; Pred. No. 1.5e-233;

Matches 693; Conservative 53; Mismatches 75; Indels 7; Gaps 6;

Qy 1 MGSFMSSEVALVQLFLPTAAAYTCVSRIGELGLVEFRDLNASVAFQRRFVVDVWRCEE 60

Db 1 MGSFMSSEVALVQLLLPTASAYNCVSQLGELGLVEFRDLNESVAFQRRFVVEVRCEE 60

Qy 61 LEKTTFTFOEVRRAVLPPPKGRLPAPPPRDLRIQEETRLAQELDRVGNQOALRA 120

Db 61 LEKTTFTFLEEVQVRAGLTTPPEGTLPAPPPRDLRIQEETRLAQELDRVGNQOALRA 120

Qy 121 QLHQLRLHSAVLGQSHSPVAAADHTEGPPSETTLLPGTRGPHSDLKUNFVAGAVEPYKA 179

Db 121 QHQLHSAVLGQSHSPVAAADHTEGPPSETTLLPGTRGPHSDLKUNFVAGAVEPYKA 180

Qy 180 PALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEGIQGKIRKITDCF 239

Db 181 AALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEGIQGKIRKITDCF 240

Qy 240 HCHVPPFLQOEERARLGALQLOQOQOEVLGETERFLS QVLGRVLQLLPPGQVQVHKM 299

Db 241 HCHVPPFLQOEERARLTLLQLOQOQOEVLGETERFLS QVLGRVQQLLPPGQVQVHKM 300

Qy 300 KAVYLALNQCSTVTHKCLIAEAWCSVRDLPALQALRDSSEEGVSAVAHRIPCRDMP 359

Db 301 KAVYLTLNQCSTVTHKCLIAEAWCAARDLPTVQALQSGSSEEGVSAVAHRIPCQDMP 360

Qy 360 TLIRNRTTASFOGIVDRYGVRYQEVNPAPTIITFPFLFAMFGDVGHGLMLFALA 419

Db 361 TLIRNRTTSSFOGIVDAVGVRYREVNPAPYTIITFPFLFAMFGDVGHGLMLFALA 420

Qy 420 MYLAENRPAVKAQNEIWTFFRGYLLMLGLFSIYTGFIYNECFSRATSIFPQSGWSVA 479

Db 421 MYLTENRPAVKAQNEIWTFFRGYLLMLGLFSIYTGFIYNECFSRATTIFPQSGWSVA 480

Qy 480 AMANQSGSDAFLAQHTMLTLDPNVTGVLPGYPFGIDPINSLAANHLSFLNSFKMKMSV 539

Db 481 AMANQSGSDEYLSQHSMLTLNPNITGVFLGYPFGIDPINSLATNHLNLSFLNSFKMKMSV 540

Qy 540 ILGVVHMAFGVVLGVFNHVFQORHLLLETLPELTFLGLFGYLVFLVYKWLVCWMAAR 599

Db 541 ILGVTHMAFGVPLSIFNHHVFQORHLLLETLPELTFLGLFGYLVFLVYKWLVSAA 600

Qy 600 AAS-PSILIHFINMFLFSQNSNRLLYPROEVQVQATLVVLALAMVPIILLGTPHLLR- 657

Db 601 ASSAPSIILHFINMFLFSQNPINLLPHQOEVOVQVVLVVLALATVPILLGTPHLLRQH 660

Qy 658 HRRRLRRRPAQOBENKAGLLDLPDASV--NGWSSDEKAGGLDDEBEAEVLVPSVLMH 714

Db 661 HRRNRTORRPAQOBEDTKLLASPDASTLNSWSPDEEKAGSGD-BETEFPVSEIPMH 719

Qy 715 QAIHTIEFCGCVSNSTASYLRWALSALAHQSEVLWAMVMRIGLGLGREGVAAVVLVP 774

Db 720 QAIHTIEFCGCVSNSTASYLRWALSALAHQSEVLWAMVMRIGLGMGREIGVAAVVLVP 779

Qy 775 IFAPAVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYSGTGKLSPP 822

Db 780 VFAAPAVLTVAILLVMEGLSAFLHALRLHWVEFQNKFSYSGTGKLSPP 827

RESULT 7

```
Q218C8 CHICK PRELIMINARY; PRT; 837 AA.
ID Q218C8 CHICK
AC Q218C8;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-FEB-2006, entry version 18.
DE Vacuolar H(+) -transporting ATPase 116 kDa subunit, a3 isoform
DE (EC 3.6.1.35).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
```

OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=white leghorn; TISSUE=Bone;
RX MEDLINE=20325091; PubMed=10866814;
RA Matteson J.P., Li X., Peng S.B., Nilsson F., Andersen P.,
RT Lundberg L.G., Stone D.K., Keeling D.J.;
RT "Properties of three isoforms of the 116-kDa subunit of vacuolar H⁺-
ATPase from a single vertebrate species. Cloning, gene expression and
RT protein characterization of functionally distinct isoforms in Gallus
RT gallus";
RL Eur. J. Biochem. 267:4115-4126(2000).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC ENBL; AJ289021; CAB93529.1; -; mRNA.
DR Ensembl; ENSGALG0000003516; Gallus gallus.
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:0020037; F-heme binding; IEA.
DR GO; GO:0016787; F-hydrolase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR012282; Cytochrome_c_r.
DR InterPro; IPR000005; HTHarac.
DR InterPro; IPR002490; V_ATPase_sub116.
DR PANTHER; PTHR11629; V_ATPase_sub116; 1.
DR Pfam; PF01496; V_ATPase_1; 1.
DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
KW Hydrolase.
SQ SEQUENCE 837 AA; 94421 MW; F5E91863B4530300 CRC64;

Query Match 62.6%; Score 2670.5; DB 2; Length 837;
Best Local Similarity 64.2%; Pred. No. 2.2e-172;
Matches 536; Conservative 85; Mismatches 195; Indels 19; Gaps 9;

QY 1 MGSFSESEVALVOLFLPTAAATCYSLGELGVFERDLNASVSAPQRFVVDVWRCEE 60
DB 1 MGSFSESEVCLAQFLHSASVCSVSELGERLLEFRDLNPHVSAPQRFVGEVRCEE 60

QY 61 LEKTFTFLOEVRAGLVLPKPPGRPLPAPPRLRIQETERLAQELDVRGNOQALRA 120
DB 61 MEKTFTFLOELHAGRVLGPCPTENPAPVAREALRVQEQELARELREVSNNALRG 120

QY 121 QLHLQHLAAVLRQGE-----POLAAHTDGASERTPLQAPGGPHODLRVNFVAGAV 174
DB 121 RLQDLRYLHLVLRGQRLTSMPPGPGSPPSRAFSSEHEPLDPSVHHLDRKINFVTGVI 180

QY 175 EPHKAPALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTPLISYWGEOIGQKIRK 234
DB 181 HPMRVNAFERLLWRACRGYLVASFVEMPEPLEDPTGTSVTWVIFLISYWGEOIGQKIRK 240

QY 235 ITDCFCHVFPFLQEQEARLGALQQLQOQSQELQEVLTGETERFLSOVLGRVQLQLPPGV 294
DB 241 ISDCFCHVFPYPPESASRADTLNGVLSQDLVSVLVEETEQLAQVLDKVLALPSWRV 300

QY 295 QVHKMAVYALNQCQSVSTTHKCLIAEAWCSVRDLPALQELALDSSMEEG--VSVAHRI 352
DB 301 QVQKMAIYLVNLQCQLDTEKCLIAEAVMCPQDLQVQELALQSGVSKSSVECFQRI 360

QY 353 PCEDMPPTLIRNRTASTQGIYDRYGVRYQVNPAPYTIITFPFLPAMFGDVGHGLL 412
DB 361 PTLESPTLIRTNKFTAGFQSIYDAYGVASYQVNPAPYAIITFPFIFAMFGDVGHGLL 420

QY 413 MELPALAMVLAENRPVAKAAQNEIMQTEFRGRVYLLLMGLFSYITGTFIYNECPSRATSLF 472
DB 421 MELPALAMVLFENSPSQ 480

QY 473 PSQSWVAAMNOSGWSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPISLANHLSFLNS 532
DB 481 PSAMSVAAMNOSGWSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPISLANHLSFLNS 540

QY 533 FRKMSVILGVVMAFGVVLGVFNHVFHQHRLLETLBELTFLGLFGYLVFLVYKW 592
DB 541 FRKMSVILGVVMAFGVVLGVFNHVFHQHRLLETLBELTFLGLFGYLVFLVYKW 600

QY 593 LCYWAARA-ASPSILIHPIFMFLFSHPNRLLYPROEVQATLVIALAMVPTLLGTP 651
DB 601 VKFSAVDSQVAPSIILHPIFMFLFSHPNRLLYPROEVQATLVIALAMVPTLLGTP 660

QY 652 LHLL-HRRRLRRRRPADROENKAGLLDLPDA--SYNGWSSDEEKAGGLDDEEAEALVP 708
DB 661 LYLKQKRRRRRANSIPPAATVEQEP-LLEGQAGNSVNATKEDVESGHHGPDAMHDF-- 717

QY 709 SEVLMHQATHITIEFCGCVSNTASYLRLWALSALHAQLSVLAMVWRIG-LGLGREVG 767
DB 718 SEVLMHQATHITIEFCGCVSNTASYLRLWALSALHAQLSVLAMVWRIG-LGLGREVG 774

QY 768 AAVLVPIPAFAVMTVAILLVMEGLSAFLHRLHWFQNFQKFSYGTGKLSPPF 822
DB 775 GGVVLPVFAFAVLTVAILLVMEGLSAFLHRLHWFQNFQKFSYGTGKLSPPF 829

RESULT 8
Q5CZ26 XENTR PRELIMINARY; PRT; 823 AA.
AC Q5CZ26;
DT 29-MAR-2005, integrated into UniProtKB/TrEMBL.
DT 29-MAR-2005, sequence version 1.
DE 07-FEB-2006, entry version 7.
DE MGC108034 protein.
GN Name=MGC108034;
OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
OX NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeng B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.C., Krzywinski M.I., Skalek U., Smalish D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole body;
RX Klein S., Gerhard D.S.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; BC090359; AAH90359.1; -; mRNA.
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR000005; HTHarac.
DR InterPro; IPR002490; V_ATPase_sub116.

```
DR Pfam; PF01496; V_ATPase_I; 1.
DR PROSITE; PS00041; HTH_Arac_FAMILY_1; UNKNOWN 1.
SQ SEQUENCE 823 AA, 93940 MW, F88783693BF5C23C CRC64;

Query Match 55.6%; Score 2370.5; DB 2; Length 823;
Best Local Similarity 56.7%; Pred. No. 5.2e-152;
Matches 476; Conservative 121; Mismatches 203; Indels 39; Gaps 10;

QY 1 MGSMFRSEVALVQLPFTAAAYTCVSRGLGELVFEFRDLNASVSFAQRRFVVDVWRCBE 60
DB 1 MGSLFRGEMCLAQFLQSGSYQCYSALGELGVEFRDLNQNVNSFORRYSEIRRCDE 60
QY 61 LEKTTFTFQEEVRRAGLVLPKGRLPAPPDRLRIQETERLAQELRDVGNQOALRA 120
DB 61 METTFSYLERLRKAGVQAPESMPPALPDRATRMQESQLAKELREVSQNTLOE 120
QY 121 QLHQQLHAAVLRLQGHGHE---POL-AAAHDTGASERTPLLOAQPFGPHQDLRVNFFVAGAVEP 176
DB 121 RLRELLEYANILRESQRFPTGLLEGEAQWKDRSEDDPLLDPAVNVKQDLRVSFMAGVHP 180
QY 177 HKAPALERLLWRACRGFLIASPRELEQPLEHPVTGEPATWMTFLISYMGEOIGQKIRKI 236
DB 181 WMNMFERLLWRACRGFLIVNFMSEPMEDIVTGESVTQIIPLISYMGKIGKIKKIA 240
QY 237 DCFHCHVFPFLQOEERLALQLOQOQSOELQEVLTGETERFLSGLVGRVQLLPPGQOVQ 296
DB 241 NCFHCHVFPYADDETSLRLETLNLLVQIQDMQVLLQTEGYSQVLSRAASALHHWRVSV 300
QY 297 HOKKAVYIALNOCVSTTHKCLIAEAWCSVRDLPALQEAIRDSSMBE--GVSAAVAHRIPC 354
DB 301 RMKHVYILNLCSV--RRCLIGEVCPVVDLPPLQLSALTRAESSGGGGSFCHRIPC 358
QY 355 RDMPTLTIRNFTASFGQIDVRYGVRQYQEVNPAPYTIITPPFLFVAFMGDVGHLMP 414
DB 359 AFSPTLTIRNFKTSFGQIDVAYGVATYQEVNPAIFSITITPPFLFVAFMGDVGHGATMF 418
QY 415 LFLALMWLAENRPVAKAQNQETWQTFFRGRYLLMLGLFSIYGTIYNCEFSRATSIIPS 474
DB 419 LFLALMWLGENDPKLRSEDEIFSCFCGRYILLMGALSVTYGFVYNECFSRPAVINS 478
QY 475 GHSVAMANQSGWSDAFLAQHTMLTLDNVTGVFLGPPYFGIDPDIWLSLAANHLSFLNSFK 534
DB 479 GHSVAMARANNTWSDSINKLPPIPLNPNITVFTAPYFGIDPDIWLSLAVNRLTFLNSFK 538
QY 535 MKMSVILGVVMAFGVLGVNHNHFGQHRLLLETLPELTLLGLRGVLPLVLYKWL 594
DB 539 MKMSVILGVCHMAFGVCLSVFNHFQRIYRITLPELLELCLFGLVPLVMVYKWI 598
QY 595 VNAARAA--SPSILIHFINMFLFSHFSNRLLYPROEVVQATLVVLALAMVPILLGLPLH 653
DB 599 LTAEDAENAPSILIHFDIMFLFTQNGRDLQGGQVQVTVLIVVAILCIPVLLGDPIC 658
QY 654 LLHRRRRLRRPADQSENKAGLLDLPDASVNGWSSDEEAKAGLDDE-----BEA 704
DB 659 LLIOHRSK-----KDSHQNRCSL-----NNGDREALLEDEITVPTGHGHS 701
QY 705 ELVPSVLMHQAHITIEFCGCVSNSTASVLRWLALSLAHQOLSEVLMAVMWRIGL-LGR 763
DB 702 KPDRAEVFMHMIHTIEFCGCVSNSTASVLRWLALSLAHQOLSEVLMAVMWRIGL 761
QY 764 EYGVAAVVLVPIFAAFVMTVAILLVMEGLSFAFLHRLHWFQNFQYSGTGKLSPPF 822
DB 762 TWG----IVLVPIFAFVAVLTVAILLMEGLSFAFLHRLHWFQNFQYSGTGKLSPPF 817

RESULT 9
Q7ZVM7 BRARE PRELIMINARY; PRT; 822 AA.
AC Q7ZVM7
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE Zgc:55891.
```

296 VHKMAYVLAALNOCVSSTHKLIAEAWCSVRDLPALQALRDSSMEEG--VSAVAHRIP 353
292 VQKCAVQTVNLCSPTDKCLIAEAWCPVSQLPALQALRGKSGNSVDFYRRLP 351
354 CRDMPPTLIRNRTASFOGIVDRYGVGRYQEVNPAVYTIITFFPFAVFGDVGHLGM 413
352 ATTSPTLPFTNSFTAGFSQSDAYGVASVYEVNPAVYTIITFFPFAVFGDVGHLGM 411
414 FLFALAMVLAENRPAVKAQNEIWTFFRGYRLLMLGLFSIYTGFIYNECFSRATIFP 473
412 TLAALMWLENDPKLRKNTNEIWRMMFGGRYLLMLGLFSIYTGFIYNECFSGKSTFS 471
474 SGHVAAMANQSGWDAFLAQHTMLTLDPNVTGVFLGPGYDIPDIPWLSLAHLSPNSP 533
472 SGHVRNPAEYFNWTESTFKKMSVLSLDPNVTGVFTGYPFGDIPDIPWGLANNHLTFLNSY 531
534 KMKMSVILGVVHMAFGVVLGVFNHVFGRHRLLETLPELTLGLGLFYLVIYKWL 593
532 KMKMSVILGVHMTFGVCLSFYFNIHPREVSVFLVLIPELCPMLCLFGLYIFWVIKWL 591
594 CWAARA-ASPSILIHFNFLSHSPSNRLLPYRQEVQATLVLAAMVPILLGTPL 652
592 VYGPVNSDSAPSILIHFDMLFTENKDKPLTYGTQVQVLAIVAVLSPVLLLGKPI 651
653 HLLHRRRRRRPADRQEBENKAGLLDLPDASVNGWSDEKAGGLDDEEAEIVPSEVL 712
652 QEYLSHKRRKRNTPGDRP-----LAEANGSINSQGDVADRGGGGGESE-BFDTANVF 704
713 MQQAIHTIEFCLGCVSNSTASYRLWALSALHAQSEVLWAMVMRIGLGLRGVGAVAL 772
705 MQQAIHTIEFCLGCVSNSTASYRLWALSALHAQSEVLWAMVMRIGLGLRGVGAVAL 762
773 VPIPAFAPVMVAIVALLMEGLSAPLHARLHWVEFQNKFGYGTGKLSPF 822
763 ALVFVGFAVLTVSILLMEGLSAPLHARLHWVEFQNKFGYGTGKLTTF 812

RESULT 10
Q4R2B2_TETNG
ID Q4R2B2_TETNG PRELIMINARY; PRT; 827 AA.
AC Q4R2B2;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Chromosome 1 SCAF14944, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTENG00026564001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouteau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaut S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Farra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX Genoscope; Whitehead Institute Centre for Genome Research;
RP

RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; CAA01014944; CAG06270.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR002490; P:proton transport; IEA.
DR PANTHER; PTHR11629; V:ATPase_sub116; 1.
DR Pfam; PF01496; V:ATPase_1; 1.
FT NON TER 827 827
SQ SEQUENCE 827 AA; 93857 MW; DE6E5AF8679C107A CRC64;

Query Match 53.8%; Score 2295.5; DB 2; Length 827;
Best Local Similarity 55.9%; Pred. No. 6.5e-147;
Matches 468; Conservative 120; Mismatches 220; Indels 29; Gaps 11;

QY 1 MGSMPRSEEEVALVQLFPTAAAYTCVSRGELGLVEFRDLNLSVAFQRRFVVDVMRCEE 60
DB 1 MGSLSRSEEEVCLVQLFLOSSAYNCVSELGELGLVEFRDLNPNVAFQRFVGEVRCEE 60
QY 61 LEKTFPLOBEVRRAGLVLPKGRLP-----APPDRLRIQEBTERLAQELRDVRG 113
DB 61 LEKTFPLOBEVRRAGLVLPKGRLP-----APPDRLRIQEBTERLAQELRDVRG 113
QY 114 NQALRAQLHQLHAALVROGHEPQLAAHTDGAERTPLQAPG--GHQDLRVNFA 171
DB 119 NRDSLRAQLTQLCOYRGVLRTH-----SITASQAPPPVLESQGLFENRDVRLSFA 172
QY 172 GAYEPHKA PALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWTMTFLISYGEQIQGK 231
DB 173 GVVHPKVASFERELLWRACRGYIIVDFREMEDLEHPEHPEHPEHPEHPEHPEHPEH 232
QY 232 IRKITCFCHVFPFLOQBEARLQALQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQO 291
DB 233 VKKICDFHTQTPAYPESQTEREELIQLQTRIEDIKSVLSQTESVLSQQLRAVAVLPQ 292
QY 292 GOVQVHKMAYVLAALNOCVSSTHKLIAEAWCSVRDLPALQALRDSSMEEG--VSAVA 349
DB 293 WKVRQCKAVQTVNLCSPTDKCLIAEAWCPVSQLPALQALRGKSGNSVDFYRRLP 352
QY 350 HRIPCRDMPPTLIRNRTASFOGIVDRYGVGRYQEVNPAVYTIITFFPFAVFGDVG 409
DB 353 NRLATSTPPTLPPLNSFTTGQNIYDAVADYREVNPVAVYTIITFFPFAVFGDVG 412
QY 410 GLMLFLFALAMVLAENRPAVKAQNEIWTFFRGYRLLMLGLFSIYTGFIYNECFSRAT 469
DB 413 GILMTLAALMWLENDPKLRKNTNEIWRMMFGGRYLLMLGLFSIYTGFIYNECFSKSL 472
QY 470 STPPSGWSVAAMANQSGWDAFLAQHTMLTLDPNVTGVFLGPGYDIPDIPWLSLAHLSP 529
DB 473 STFNSGHVKRPMDDNNVNSVLSGTQFLPMDVPVPGFTSPYFPGDIPWGLANNKLT 532
QY 530 LNSFKMKMSVILGVVHMAFGVVLGVFNHVFGRHRLLETLPELTLGLGLFYLVIYKWL 589
DB 533 LNSYKMKMSVILGVHMTFGVCLSFYFNIHPREVSVFLVLIPELCPMLCLFGLYIFW 592
QY 590 YKWLVCWAARA-SPSILIHFNFLSHSPSNRLLPYRQEVQATLVLAAMVPILLGTPL 648
DB 593 FKWIVVTPAQSKPAPSILIHFDMLFTENKDKPLTYGTQVQVLAIVAVLSPVLLLGKPI 652
QY 649 GTPL--HLLHRRRR--FLRRRPADRQEBENKAGLLDLPDASVNGWSDEEAKAGGLDDEEAE 705
DB 653 GKPICKYVYTKKRRSHVTSCKLQEDRHLPMG--DSSNINTSRSGVEEG--STENEAV 708
QY 706 LVPSEVLMEQAIHTIEFCLGCVSNSTASYRLWALSALHAQSEVLWAMVMRIGLGLRGV 765
DB 709 FDTADVLMQAIHTIEFCLGCVSNSTASYRLWALSALHAQSEVLWAMVMRIGLGLRGV 768

766 GAAAVLVPIFAFAVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPF 822
 769 G--AAVLVPIFAFAVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPF 823

RESULT 11
 QARS23 TETNG
 ID QARS23 TETNG PRELIMINARY; PRT; 838 AA.
 AC QARS23 TETNG
 DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
 DT 19-JUL-2005, sequence version 1.
 DT 07-FEB-2006, entry version 4.
 DE Chromosome 12 SCAP14999, whole genome shotgun sequence. (Fragment).
 GN ORFNames=GSTENG0029489001;
 OS Tetraodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthopterygii; Acanthopterygii; Perciformes; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 OX NCBI_TaxID=99883;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX PubMed=15496914; DOI=10.1038/nature03025;
 RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
 RA Mauceli E., Bounneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segreus B.,
 RA Desilva C., Salancoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
 RA Bionnet C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
 RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
 RA Kellis M., Volff J.-N., Gulgo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
 RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,
 RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 RT the early vertebrate proto-karyotype.";
 RL Nature 431:946-957(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope, Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -! CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 CC EMBL; CAAG0101499; CAG08489.1; -; Genomic DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0020037; F:heme binding; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0015992; P:proton transport; IEA.
 DR InterPro; IPR002355; Cu oxidase Cu BS.
 DR InterPro; IPR002490; V:ATPase sub16.
 DR PANTHER; PTHR11629; V:ATPase sub116; 1.
 DR Pfam; PF01496; V:ATPase_1; 1.
 DR PROSITE; PS00079; MULTICOPPER_OXIDASE1; UNKNOWN_1.
 DR NON_TER 838
 SQ SEQUENCE 838 AA; 96100 MW; 5B871FC22DA381EA CRC64;

Query Match
 Best Local Similarity 51.3%; Score 2188; DB 2; Length 838;
 Matches 437; Conservative 130; Mismatches 247; Indels 20; Gaps 8;

1 MGSMFRSEVALVQLFLPTAAAYTCVSRIGELGLVEFRDLNLSAFAQRFFVVDVWRCCE 60
 1 MGSLFRGEICLAQLFLQAGSAYDCISGELGLVEFRDLNPTVTFQRYKYSKICEE 60

61 LEKTFPQEEVRRAGLVLPKPKGLPAPPDRLRIQETRLAQELRDVGNQOALRA 120
 61 MERILGYLMKEVKKADISLPEGDVNPPIAPLPKHILSIMEQLQRLVELGEVTKNKEQLR 120

121 QL-----HQQLHAAVLRQGHREPOLAAAHDTGAS--ERTPELLQAPGPHQDLRVNFA 171
 121 NLLLETEYHMLRITRNFVQSAERENALHREFFLEKDTMDYSSMRIGLAKGLPFVS 180

172 GAVEPHKAPALERLLWRACRGFLIASFRELEOPLEHPVTGBPATWMTFLISVWGEIQOK 231
 181 GIIQVKIEAFERMLRWCKYITLTYAEVYLENPDGTGETKSVVFLISVWGEIQOK 240

232 IRKITDCPHCHVFPFLQOEERLALQQLQOQSOBELQEVLGETERPLSQVLGRVLQLPP 291
 241 VKKICDCYCHLYPYSPSSNEERNDVLEGLKTRIQDLHTVLHRTEDYLRQVLVKASSEIYT 300

292 GOVOYHKKAVYLAALNOCSSVTHKCLTAECVSRDIPALQEARLDSMEG--VSAVA 349
 301 WTIQVKMKAIYIILNLCSDVTNKCLTAECVSRDIPALQEARLDSMEG--VSAVA 360

350 HRIPCRDMPPTLIRNTRFTASFGQIVDRVGVQYOVNAPVYTIITFPFLFAMVPGDVGH 409
 361 NRIPTSSPTLIRNTRFTASFGQIVDRVGVQYOVNAPVYTIITFPFLFAMVPGDVGH 420

410 GLLMFLFALAMVLAENRPAVKAQNEIWQTFPRGRYLLMLGLFSIYTGFIYNECFSRAT 469
 421 GLIMLFAMVLYENNRKVKXNRNEIWNMPPEGRYIILMLGLFSIYTGFIYNECFSRAT 480

470 SIPTSGMSVAMANOSGSDAFLOHTMTLDPNTVGTGFLGYPFGIDPINSIAHLSPF 529
 481 NIFSGMSVAMFRENVMQDDVYGNRFLTDPNTVGTGFLGYPFGIDPINSIAHLSPF 540

530 LNSPFMKMSVILGVVHMAFGVVLGVFNHVFQQRHLLLETPELTGLGLGYLVFLVI 589
 541 LNSYFKMSVILGVVHMAFGVVLGVFNHVFQQRHLLLETPELTGLGLGYLVFLVI 600

590 YKWLVCWAAARA--SPSIIHFINMPLFSHSPSNRLLYPRQEVVQATLVVLAAMVPILL 648
 601 YKWLAFSAKDSRHAPSILHFINMPLMQGS--AMQPLYPQNGQLFVAVVAVLSVPVFL 659

649 GTPHLHLHRRRLRRRPAQRQENKAGLLDLPDASVNGWSDEKAGGLDDEEAEVLP 708
 660 GRPLLYLWLNHNGNPFSTCGIGDEE--LFLLRADDMEGSSSHSDPSSSG---DHQSENFN 714

709 SEVLHQHAIHTIEFCLGCVSNSTASYLRWLALSHLAHQSEVLWAMVMRTGLGLREVGA 768
 715 ADELHQHAIHTIEFCLGCVSNSTASYLRWLALSHLAHQSEVLWAMVMRTGLGLREVGA 773

769 AVLVPIFAFAVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPF 822
 774 -LFLVPIFAFAVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPF 826

RESULT 12
 VPP2 BOVIN
 ID VPP2 BOVIN STANDARD; PRT; 854 AA.
 AC Q97681;
 DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
 DT 01-MAY-1999, sequence version 1.
 DT 07-FEB-2006, entry version 34.
 DE Vacuolar proton translocating ATPase 116 kDa subunit A isoform 2 (V-ATPase 116-kDa isoform a2).
 GN Name=ATP6V0A2; Synonyms=ATP6N1B;
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 OC Pecora; Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA].
 RX MEDLINE=99107917; PubMed=9891027; DOI=10.1074/jbc.274.4.2549;
 RA Peng S.-B., Li X., Crider B.P., Zhou Z., Andersen P., Tsai S.J.,
 RA Xie X.-S., Stone D.K.;
 RT "Identical and reconstitution of an isoform of the 116-kDa
 RT subunit of the vacuolar proton translocating ATPase.";
 RL J. Biol. Chem. 274:2549-2555(1999).

CC -!- FUNCTION: Part of the proton channel of V-ATPases.
CC -!- SUBUNIT: The V-ATPase is an heteromultimeric enzyme composed of at
CC least thirteen different subunits. It has a membrane peripheral V1
CC sector for ATP hydrolysis and an integral V0 for proton
CC translocation. The V1 sector comprises subunits A-H, whereas V0
CC includes subunits a, d, c, c', and c''.
CC -!- SUBCELLULAR LOCATION: Membrane; multi-pass membrane protein.
CC -!- TISSUE SPECIFICITY: Highly expressed in lung, kidney and spleen.
CC -!- SIMILARITY: Belongs to the V-ATPase 116 kDa subunit family.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL: AF105016; AA012058.1; ; mRNA.
CC InterPro: IPR002490; V-ATPase_sub116.
CC PANTHER: PTHR11629; V-ATPase_sub116; 1.
CC Pfam: PF01496; V-ATPase_I; 1.
CC Glycoprotein; Hydrogen ion transport; Ion transport; Membrane;
CC Transmembrane; Transport.
CC
CC CHAIN 1 854 Vacuolar proton translocating ATPase 116
CC kDa subunit A isoform 2.
CC /FTID=PRO_0000119215.
CC Extracellular (Potential).
CC
CC TOPO_DOM 1 396 Potential.
CC TRANSMEM 397 413 Potential.
CC TRANSMEM 452 472 Potential.
CC TRANSMEM 555 575 Potential.
CC TRANSMEM 591 611 Potential.
CC TRANSMEM 651 671 Potential.
CC TRANSMEM 787 807 Potential.
CC CARBOHYD 43 43 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 157 157 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 438 438 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 484 484 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 505 505 N-linked (GlcNAc...) (Potential).
CC
CC SQ SEQUENCE 854 AA; 98010 MW; 8BD9A128465CCED5 CRC64;
CC
CC Query Match 50.48; Score 2149; DB 1; Length 854;
CC Best Local Similarity 50.68; Pred. No. 6e-137;
CC Matches 439; Conservative 133; Mismatches 222; Indels 74; Gaps 15;
CC
CC QY 1 MGSFMRSEAEVALVQLFLPTAAATCVSRGELGIVEPRDLNASVAFQRFVVDVWRCCE 60
CC DB 1 MGSLSRSETNCLAQFLQSTAVECLSVLGEKGLVEPRDLNQNVSSQRFVGVKRCCE 60
CC
CC QY 61 LEKTFPQBEVRRAGVLVPPKGRLPAPPRLRLRIQETTERLAQELRDVRGNQALRA 120
CC DB 61 LERILAYLQVEINRADIPLEGDTSPAPPLKQVLEMOEQLOKLELREVTNKKELRK 120
CC
CC QY 121 QL-----HQLQLHAALVLRQGE-----PQLAAAHDTGASERTPLLOAQPQGHQDL 165
CC DB 121 NLELEIYTHMLRVTKTFVKNVVEFETPEYEEFPPL-----ENESLLDYSQMQLGA 171
CC
CC QY 166 RVNFVAGAVPHKAPALERLLWRACRGFLIASFRELEQPLEHPVTGSPATMTFLISYWG 225
CC DB 172 KLGFVSLINGQKVEAFKMLRWCKGYTIVTAELDEPLEDETGETGEVVKWYVFLISFWG 231
CC
CC QY 226 EQIGQKIRKITDCHGHVFPFLQOEAREALGALQLOOQSOELQVLETERFLSLQVLRV 285
CC DB 232 EQIGHVKYKICDCYCHVYYPNTABERREIQGLNTRIQLDLYTLVHKETDYLRLVQLCA 291
CC
CC QY 286 LQLLPPQGVQVHKMKAVYALNOCVSTTHKLIAGWCSVRDLPAQELALRDSMEEG- 344
CC DB 292 AESVYGRVIVQKKKAIYHMLNMCSDFTVNKLIAGWCSVRDLPAQELALRDSMEEG- 351
CC
CC QY 345 -VSAAHRIPCRDPPTLIRNFTASFGQIVDRYGVGRYQEVNPAFTYITIPFPLFAVM 403
CC DB 352 TIFSFMTNPTKETPTPLIRNFTASFGQIVDRYGVGRYQEVNPAFTYITIPFPLFAVM 411
CC
CC QY 404 FGVGVHGLMFLPALAMVLAENPAVKAQNEIWTQTFRGRYLLMLGLFSYITGYTNE 463
CC DB 412 FGDGFGHGVNMFALLLVLENHPRNLQSQ-EIWRMFFNCRGYILLMLGLFSYITGLYND 470
CC
CC QY 464 CFSRATSI PPSGSHVAAANQSG-----WSDAFLAQHTMLTLDPNVTGVFLGYPY 513

Db 471 CFSKSVNLFGRSNVNSAMYSSSHSPBQRKMWLWDSIVRHHSVLQLDPSVPGFRGYP 530
QY 514 FGIDPIWLSAANHLFLNSFKMKSIVLGVVHMAFGVVLGVVFNHVFQGRHRLLETLPE 573
Db 531 FGIDPIWNLATNRLFLNSFKMKSIVLGVVHMAFGVVLGVVFNHVFQGRHRLLETLPE 590
QY 574 LTFLLGFGYLVFLVIYKVLGVMAAA--SPSILIHFMFLFSHSPSRLLYPROEVV 631
Db 591 LLEMLCIFGLYLIIFMIYKWL-VYSAETSRTPSILIEFISWFLFLASDTGG-LYPGQEHV 648
QY 632 QATLVVLAMVPIILLGLTPLLHLHRRR-----LRRRPADROENKAGLLDL 680
Db 649 QRLLLITLVLSVPLFLGKPLFLWLHGRSCFVGSGYTLVRK--DSEEE----- 698
QY 681 PDASVNGWSDEEKAGGLDDE-----EEAELVPSEVLMEQAHITIEFCIGCVSNTASYL 734
Db 699 --VSLGGQDIEEGNQMEDGCEVTCEDFD--GEILMTQIHSIYCLGICISNTASYL 754
QY 735 RLWALSIAHQAQLSEVLWAMVNRIGLGRGVAAVVLVPIFAFAVMTVAILLVMEGLS 794
Db 755 RLWALSIAHQAQLSEVLWAMVNRIGLGRGVAAVVLVPIFAFAVMTVAILLVMEGLS 812
QY 795 AFLHALRLHWMVFQNKFSYSGYKLSPP 822
Db 813 AFLHALRLHWMVFQNKFSYSGYKLSPP 840
RESULT 13
VPP2_MOUSE
ID_VPP2_MOUSE STANDARD; PRT; 856 AA.
AC P15920; Q3UX3; Q9UJ02;
DT 01-APR-1990, integrated into UniProtKB/Swiss-Prot.
DT 11-JAN-2001, sequence version 2.
DT 07-FEB-2006, entry version 57.
DE Vacuolar proton translocating ATPase 116 kDa subunit A isoform 2 (V-
DE ATPase 116-kDa isoform a2) (immune suppressor factor J657).
GN Name=AtP60a2; Synonyms=AtP6n1b, Tj6;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RX MEDLINE=91061805; PubMed=2247090; DOI=10.1016/0161-5890(90)90102-6;
RA Lee C.-K., Ghoshal K., Beaman K.D.;
RT "Cloning of a cDNA for a T cell produced molecule with a putative
RT immune regulatory role.";
RL Mol. Immunol. 27:1137-1144 (1990).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA].
RX MEDLINE=20187595; PubMed=10722719; DOI=10.1074/jbc.275.12.8760;
RA Toyonura T., Oka T., Yamaguchi C., Wada Y., Futai M.;
RT "Three subunit A isoforms of mouse vacuolar H+-ATPase. Preferential
RT expression of the a3 isoform during osteoclast differentiation.";
RL J. Biol. Chem. 275:8760-8765 (2000).
RN [3]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Brain, and Heart;
RX MEDLINE=20167151; PubMed=10702241; DOI=10.1074/jbc.275.10.6824;
RA Nishi T., Forgac M.;
RT "Molecular cloning and expression of three isoforms of the 100-kDa a
RT subunit of the mouse vacuolar proton-translocating ATPase.";
RL J. Biol. Chem. 275:6824-6830 (2000).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=NOD; TISSUE=Dendritic cell;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,

RA Ambesi-Impionbato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
 RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
 RA Guetlich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
 RA Hall D., Huminieski L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
 RA Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 RA Mottagui-Taber S., Mulder N., Nakano N., Nakachi H., Ng P.,
 RA Nakazoni R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavesi G., Pesole G.,
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 RA Roest B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
 RA Schonbach C., Sekiguchi K., Sempile C.A., Seno S., Sessa L., Sheng Y.,
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 RA Sperling S., Stupka E., Sugiyara K., Sultana R., Takenaka Y., Taki K.,
 RA Tamajima K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
 RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
 RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashita N.,
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
 RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
 RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
 RA Hayashizaki Y.,
 RA "The transcriptional landscape of the mammalian genome";
 RT Science 309:1559-1563 (2005).
 RN [5]
 RN NCLEUTIDE SEQUENCE [LARGE SCALE MRNA].
 RP MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strauber R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.F., Shuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.P., Schetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abranson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X.J., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting J.M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buttsfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences".
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -I- FUNCTION: Part of the proton channel of V-ATPases (By similarity).
 CC Seems to be involved in T cell suppression.
 CC -I- SUBUNIT: The V-ATPase is an heteromultimeric enzyme composed of at
 CC least thirteen different subunits. It has a membrane peripheral V1
 CC sector for ATP hydrolysis and an integral V0 for proton
 CC translocation. The V1 sector comprises subunits A-H, whereas V0
 CC includes subunits a, d, c, c', and c'',
 CC -I- SUBCELLULAR LOCATION: Membrane; multi-pass membrane protein (By
 CC similarity).
 CC -I- TISSUE SPECIFICITY: Thymus.
 CC -I- SIMILARITY: Belongs to the V-ATPase 116 kDa subunit family.
 CC
 CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
 CC Distributed under the Creative Commons Attribution-NoDerivs License

| |
|---|
| EMBL; M31226; AAA93336.1; -; mRNA. |
| DR EMBL; X55184; CAA38968.1; -; mRNA. |
| DR EMBL; AB022323; BAA93007.1; -; mRNA. |
| DR EMBL; AF218252; AAF59921.1; -; mRNA. |
| DR EMBL; AK155055; BAE33017.1; -; mRNA. |
| DR EMBL; BC108991; AA108992.1; -; mRNA. |
| DR EMBL; BC108992; AA108993.1; -; mRNA. |
| DR PIR; JH0287; JH0287. |
| DR Ensembl; ENSMUSG0000038023; Mus musculus. |
| DR MGI; MGI.104855; Atg6v0a2. |
| DR GO; GO:0001669; C:cytosome; IDA. |
| DR GO; GO:0005615; C:extracellular space; TAS. |
| DR GO; GO:0016021; C:integral to membrane; TAS. |
| DR InterPro; IPR002490; V ATPase sub116. |
| DR Pfam; PF01436; V ATPase I; 1. |
| KW Glycoprotein; Hydrogen ion transport; Ion transport; Membrane; |
| KW Transmembrane; Transport. |
| FT CHAIN 1 856 |
| FT FT |
| FT TOPO_DOM 1 396 |
| FT TRANSMEM 397 413 |
| FT TRANSMEM 452 472 |
| FT TRANSMEM 555 575 |
| FT TRANSMEM 586 606 |
| FT TRANSMEM 651 671 |
| FT TRANSMEM 787 807 |
| FT CARBOHYD 43 43 |
| FT CARBOHYD 157 157 |
| FT CARBOHYD 505 505 |
| FT CONFLICT 486 486 |
| FT CONFLICT 791 791 |
| FT SEQUENCE 856 AA; 98145 MW; 6A0D593F6F401E22 CRC64; |
| Query Match 50.3%; Score 2146; DB 1; Length 856; |
| Best Local Similarity 51.4%; Pred. No. 9.6e-137; |
| Matches 441; Conservative 131; Mismatches 232; Indels 54; Gaps 151; |
| QY 1 MGSMFRSEVALVOLFLPTAAATCTVSRIGELGLVEPRDLNASVSARQRFFVDVWRCEE 60 |
| DB 1 MGSLFRSSECLAQFLQSQTAYECLSALGKGLVQLPDRNQNVSPQRKFVGSKRCEE 60 |
| QY 61 LEXTFTPLQEVRRAGLVLPPPGRPLPAPPRLDLLRIQEETERLAQELRDVRGNQQALRA 120 |
| DB 61 LERILLYLVEITRADIPLEGASPPAPPLKHVLEMQEQKLKEVELREVTKNKEKLKLR 120 |
| QY 121 QL-----HQLOIHAAVLRQGHE-----POLAAAHDTGASERTPLLOAPGGPHQDLRVN 168 |
| DB 121 NLDELVEYTHMLRVTKTFLKRNVFEPTYEEFPALENDSLDDYS-CMQLGA-----KLG 174 |
| QY 169 FVAGAVEPKAPALELLHWACRGFLIASPRELEOPLHPVTGEPAWTMTFLISYWGEQI 228 |
| DB 175 FVSGLIQQGRVEAFERMLMRACKGYITVTVAELDDECLDPETGEVIKWYVFLISFWGEQI 234 |
| QY 229 GQIRKITDCFHCVPPFLQOEAREALCALQOOSQELCEVLGETERTFSQVLGRVLQL 288 |
| DB 235 GHVKVKICDYCHCIYPNTAEREERI QEGLNTRIQDLTVLHKTEDYLROVLCKAAES 294 |
| QY 289 LPFGQOVQHMKAVYIALNOCSSYSTHKCLIAABWCSDRLPALQEAALRDSSMBEEG--VS 346 |
| DB 295 VCSRVVQVRKMKAIFYHLNMKCPDVNTKCLIAEAWCPEVDLPGLRRALIEGSRSSGATIP 354 |
| QY 347 AVAHRIRPCRDMPPLLRTNRFTFASFGQIVDRYGVGRYQEVNPAPYTIITTPPFLFAVMFGD 406 |
| DB 355 SFNNTIPTKETPTPLIRTNNKTTFGFGQNI VDAYGVGSYREYNPALFTIITTPPFLFAVMFGD 414 |
| QY 407 VGHLGLMFLPALAMWLAEENRPVAKAAQNEITWQTFFRGRYLLLMGLFSITYGTINYSCFS 466 |
| DB 415 FHGHGFVWFLLALLVLENHPIRSQSQ-EILLRNPFDFGRYILLMLGLEFSVVTGLIYNDCFS 473 |
| QY 467 RATSIFPGSGWSVAAMNQSG-----WSDAFLAQHTMLTLDPNVTGVGLGPYFPGI 516 |

```
Db 474 KSNVLFSGWNVVSAMYSSSHSPERQKMLWNSDSTIRHSTRLQDLPNIPGVFRGYPYFPGI 533
Qy 517 DPTWSLAANHLSPNSFKMKSIVLGVVHAFVGVVFNHVFQGRHRLLELTPBLTF 576
Db 534 DPTWNLATNRLTFLNSFKMKSIVLGFHMTFGVVLGFHNLHPRKFNVLVSVPEILF 593
Qy 577 LLGLFGYLVFLVIYKMLCVWAARA-ASPSILIHFINNLFSSHSPSNRLLYPRQEVQATL 635
Db 594 MLCIFGYLIIFMIYIKWLAYSAAETSRAPSILIEFINNLPFTSKTHG-LYPGQAHVQVRL 652
Qy 636 VVLALAMVPILLGLTPLHLLHRRR-----LRRPADROENKAGLLDLPDAS 684
Db 653 VALTVLAVPVFLGKPLFLWLHNGRNCFCGSRSGYTLVRK--DSEEE--VSLIGNODIE 708
Qy 685 VNGWSSDEEKAGGLDDEEAEELV-----SEVLMOHAIHTIEFCGVCVSNNTASYLRLWAL 744
Db 709 -EGNSRME--GCREVTCEBFEFGEILMTQAIHSIEYCLGICISNTASYLRLWALSAAH 764
Qy 745 QLSVLMWVMRIGLGLGREGVAAVVLVPIFAFAVMTVAILLVMEGLSAFLHALRH 804
Db 765 QLSVDVLMWMLRVGLRVDTTYGV--LLLLPVMFAFVLTIFILLVMEGLSAFLHAIRLHW 822
Qy 805 VEFONKFSYSGTYKLSPP 822
Db 823 VEFQNKFYVAGTKFVPF 840

RESULT 14
Q216B1_RAT
ID Q216B1_RAT PRELIMINARY; PRT; 856 AA.
AC Q216B1
DT 07-MAR-2006, integrated into UniProtKB/TrEMBL.
DT 07-MAR-2006, sequence version 1.
DT V-H+ATPase subunit a2.
DE Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Wistar;
RA Guyon S., Amar M., Fossier P., Morel N.;
RT "Neurons coexpress three different v-ATPase subunit a isoforms.";
RL Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; DQ286425; ABB91444.1; -; mRNA.
SQ SEQUENCE 856 AA; 98041 MW; B2180D770C133469 CRC64;

Query Match 50.3%; Score 2145; DB 2; Length 856;
Best Local Similarity 50.5%; Pred. No. 1.le-136;
Matches 436; Conservative 137; Mismatches 226; Indels 64; Gaps 15;

Qy 1 MGSMFRSEVALVQLFLPTAAATCVSRIGELGLVEPRDLNLSVAFQRFVVDVNRCEE 60
Db 1 MGSILFSESNCLAQFLQSTATVECLSGALKEGLVQFRDNLQNVSSQRFVGEVNRCEE 60
Qy 61 LEKFTTFLQEVRRAGLVLPKGRLPAPPRLRLRIQEETERLAQELRDVRGNQQAALRA 120
Db 61 LERILVYLQEIETRADIPLEGEASPPAPPVKHVLEMQEQLKLELRELVRTNKKELKRL 120
Qy 121 QL-----HQLQHAVALRQGE-----PQLAAHTDGSERTPPLQAPGGPHQDLRVN 168
Db 121 NLLELVEYTHMLRVTKTFLKRVNPFEPTEYEFPALESDSLDDYS-CMQRILGA-----KLG 174
Qy 169 FVAGAVEPHKAPALERLLWRACRGFLIASPRELEQPLEHPVTGEPATMTFLISYWEQEI 228
Db 175 FVSGELTQQKVEAFERMLWRACKGYITVTAELDEADELDEPTEGVIKWYVFLISFWEQEI 234
```

```
Qy 229 GQKIRKTTDFCHVFPFLOEQEARLGALQLOOQSQEOLQEVLGETERFPLSQVLGRVLQL 288
Db 235 GHVKVKICDCVCHYIYPNPTABERREIQELNTRIQLVTLVHKTEDYLKQLCKAAS 294
Qy 289 LPQGVQVQKWKAVYALNOCVSTTHKCLIAEAWCSVRDLPAEQEALRDSMEEG--VS 346
Db 295 VCSRVQVRKWKAYIYHMLNMCSPDVINKCLIAEAWCEVDLPGLRRALESGRESGATIP 354
Qy 347 AVAHRIPCRDMPPTLIRTNRTASFOGIVDRYGVGRYQEVNPAPTYITITPPFLFAVFGD 406
Db 355 SFMTIPTKETPTLIRTNKFTGQFQNVDAVGVSREVNPALFTIITPPFLFGVMFGD 414
Qy 407 VGHGLMFLPALAMVLAENRPVAKAAQNEIWTQFFRGYRLLMLGLFSITGFTYNECPFS 466
Db 415 LGHGFVFLFALLVLNENHPRLSQS-Q-EILRMFFDGRYILLMLGLFSVTGLTYNDCFS 473
Qy 467 RATSIPSGWSVAAMANOSG-----WSDAFLAQHTMLTLDPNVTGVFLGYPYFPGI 516
Db 474 KSNLFGSGWNVVSAMYSSSHSPERQKMLWNSDSTIRHSTRLQDLPNIPGVFRGYPYFPGI 533
Qy 517 DPTWSLAANHLSPNSFKMKSIVLGVVHAFVGVVFNHVFQGRHRLLELTPBLTF 576
Db 534 DPTWNLATNRLTFLNSFKMKSIVLGFHMTFGVVLGFHNLHPRKFNVLVSVPEILF 593
Qy 577 LLGLFGYLVFLVIYKMLCVWAARA-ASPSILIHFINNLFSSHSPSNRLLYPRQEVQATL 635
Db 594 MLCIFGYLIIFMIYIKWLAYSAAETSRAPSILIEFINNLPFTSKTHG-LYPGQAHVQVRL 652
Qy 636 VVLALAMVPILLGLTPLHLLHRRR-----LRRPADROENKAGLLDLPDAS 684
Db 653 VALTVLAVPVFLGKPLFLWLHNGRSCFGVSRSGYTLVRK--DSEEE-----VS 700
Qy 685 VNGWSSDEEKAGGLDDEEAEELV-----SEVLMOHAIHTIEFCGVCVSNNTASYLRLWAL 739
Db 701 LLG-SQDIEEGNNRMEEGCREMTCEBFEFGEILMTQAIHSIEYCLGICISNTASYLRLWAL 759
Qy 740 SLAHAQLSEVLWAMVRIGLGLGREGVAAVVLVPIFAFAVMTVAILLVMEGLSAFLHA 799
Db 760 SLAHAQLSDVLMWMLRVGLRVDTTYGV--LLLLPVMFAFVLTIFILLVMEGLSAFLHA 817
Qy 800 LRLHWWFEONKFSYSGTYKLSPP 822
Db 818 IRLHWWFEONKFSYVAGTKFVPF 840

RESULT 15
Q4S789_TETNG
ID Q4S789_TETNG PRELIMINARY; PRT; 897 AA.
AC Q4S789;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Chromosome 4 SCAF14575, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTENG00017316001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jaillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer S., Lutfalla G., Dossat C., Segurens B.,
RA Nicaut C., Jaffe D., Fisher S., Lutfalla G., Boudet N., Castellano S.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Kattinka M., Vacherie B.,
RA Anthouard V., Jubin C., Castelli V., Kattinka M., Vacherie B.,
RA Blemont C., Skalli Z., Cattolico L., Poullain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
```


GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2006, 13:08:24 ; Search time 52 Seconds
(without alignments)
1383.657 Million cell updates/sec

Title: US-10-783-519-2
Perfect score: 4266
Sequence: 1 MGSMPSEEEVALVQLFLPTA.....HWVEFQNKFSYGTGKLSKPF 822

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /EMC Celerra_SID33/ptodata/2/iaa/5 COMB.pep.*
- 2: /EMC Celerra_SID33/ptodata/2/iaa/6 COMB.pep.*
- 3: /EMC Celerra_SID33/ptodata/2/iaa/7 COMB.pep.*
- 4: /EMC Celerra_SID33/ptodata/2/iaa/H COMB.pep.*
- 5: /EMC Celerra_SID33/ptodata/2/iaa/ECTUS COMB.pep.*
- 6: /EMC Celerra_SID33/ptodata/2/iaa/RE COMB.pep.*
- 7: /EMC Celerra_SID33/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|----------------------|
| 1 | 4266 | 100.0 | 822 | 2 | US-08-684-932A-38 |
| 2 | 4266 | 100.0 | 822 | 2 | US-09-618-304B-2 |
| 3 | 4235.5 | 99.3 | 849 | 2 | US-09-949-016-8846 |
| 4 | 4235.5 | 99.3 | 849 | 2 | US-09-949-016-8847 |
| 5 | 4235.5 | 99.3 | 853 | 2 | US-09-949-016-8275 |
| 6 | 4235.5 | 99.3 | 853 | 2 | US-09-949-016-8276 |
| 7 | 3394 | 79.6 | 750 | 2 | US-09-949-016-7201 |
| 8 | 3394 | 79.6 | 750 | 2 | US-09-949-016-7202 |
| 9 | 2143.5 | 50.2 | 855 | 7 | 519526-1 |
| 10 | 1801.5 | 42.2 | 847 | 2 | US-09-618-304B-3 |
| 11 | 1097.5 | 25.7 | 513 | 2 | US-10-094-749-2211 |
| 12 | 892.5 | 20.9 | 515 | 2 | US-09-270-767-43227 |
| 13 | 757.5 | 17.8 | 386 | 2 | US-09-248-796A-17451 |
| 14 | 704 | 16.5 | 380 | 2 | US-09-270-767-44518 |
| 15 | 513 | 12.0 | 257 | 2 | US-09-270-767-33442 |
| 16 | 513 | 12.0 | 257 | 2 | US-09-270-767-48359 |
| 17 | 478 | 11.2 | 343 | 2 | US-09-270-767-45290 |
| 18 | 427 | 10.0 | 163 | 2 | US-09-270-767-43040 |
| 19 | 418.5 | 9.8 | 425 | 2 | US-09-248-796A-17452 |
| 20 | 370 | 8.7 | 155 | 2 | US-09-270-767-31642 |
| 21 | 370 | 8.7 | 155 | 2 | US-09-270-767-46859 |
| 22 | 366 | 8.6 | 196 | 2 | US-09-248-796A-17445 |
| 23 | 319.5 | 7.5 | 154 | 2 | US-09-270-767-32407 |
| 24 | 319.5 | 7.5 | 154 | 2 | US-09-270-767-47624 |
| 25 | 278 | 6.5 | 662 | 2 | US-09-134-000C-5682 |
| 26 | 276 | 6.5 | 186 | 2 | US-09-248-796A-17453 |

| | | | | | | |
|----|-------|-----|------|---|----------------------|-------------------|
| 27 | 226.5 | 5.3 | 359 | 2 | US-09-248-796A-17449 | Sequence 17449, A |
| 28 | 210 | 4.9 | 129 | 2 | US-09-270-767-58569 | Sequence 58569, A |
| 29 | 197 | 4.6 | 171 | 2 | US-09-270-767-32967 | Sequence 32967, A |
| 30 | 163 | 3.8 | 32 | 2 | US-09-618-304B-4 | Sequence 4, Appli |
| 31 | 162 | 3.8 | 30 | 2 | US-09-618-304B-7 | Sequence 7, Appli |
| 32 | 155.5 | 3.6 | 660 | 2 | US-09-198-452A-109 | Sequence 109, App |
| 33 | 154.5 | 3.6 | 650 | 2 | US-09-438-185A-93 | Sequence 93, Appl |
| 34 | 150 | 3.5 | 33 | 2 | US-09-618-304B-9 | Sequence 9, Appli |
| 35 | 138.5 | 3.2 | 1165 | 2 | US-09-949-016-11392 | Sequence 11392, A |
| 36 | 137.5 | 3.2 | 1165 | 2 | US-09-949-016-6874 | Sequence 6874, Ap |
| 37 | 136 | 3.2 | 26 | 2 | US-09-618-304B-5 | Sequence 5, Appli |
| 38 | 131 | 3.1 | 1985 | 2 | US-09-495-714C-6 | Sequence 6, Appli |
| 39 | 124.5 | 2.9 | 1912 | 2 | US-09-495-714C-2 | Sequence 2, Appli |
| 40 | 124.5 | 2.9 | 1977 | 2 | US-09-495-714C-4 | Sequence 4, Appli |
| 41 | 117.5 | 2.8 | 547 | 2 | US-09-252-991A-27640 | Sequence 27640, A |
| 42 | 116.5 | 2.7 | 178 | 2 | US-09-348-796A-17450 | Sequence 17450, A |
| 43 | 115 | 2.7 | 703 | 2 | US-09-252-991A-18391 | Sequence 18391, A |
| 44 | 114.5 | 2.7 | 621 | 2 | US-09-720-317A-16 | Sequence 16, Appl |
| 45 | 114.5 | 2.7 | 782 | 2 | US-09-949-016-6943 | Sequence 6943, Ap |

ALIGNMENTS

RESULT 1
US-08-684-932A-38
; Sequence 38, Application US/08684932A
; Patent No. 6403304
; GENERAL INFORMATION:
; APPLICANT: Stashenko, Philip
; APPLICANT: Li, Yi-Ping
; APPLICANT: Wucherpfennig, Anne L.
; TITLE OF INVENTION: HUMAN OSTEOCLAST-SPECIFIC AND -RELATED
; TITLE OF INVENTION: DNA SEQUENCES
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/684,932A
; FILING DATE: 19-JUL-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: FDC92-02FM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-684-932A-38

Query Match 100.0%; Score 4266; DB 2; Length 822;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSMPSEEEVALVQLFLPTAAATCVSRUGELGLVEFRDLNASVSAPQRFVVDVWRCEE 60
|||||
DB 1 MGSMPSEEEVALVQLFLPTAAATCVSRUGELGLVEFRDLNASVSAPQRFVVDVWRCEE 60

61 LEKTTFFLOEVRRAAGLVLPKGRLLPAPPPDRLRIQEETRLAQELRDVRGNOALRA 120
61 LEKTTFFLOEVRRAAGLVLPKGRLLPAPPPDRLRIQEETRLAQELRDVRGNOALRA 120
121 QHQQLHAAVLRQGHPEPQAAAAHTDASERTPLQAQGGPHQDLRVNFVAGAVPHKAP 180
121 QHQQLHAAVLRQGHPEPQAAAAHTDASERTPLQAQGGPHQDLRVNFVAGAVPHKAP 180
181 ALERLLWRACRGFLIASPRELEQPLEHPVTGEPATWMTFLISYWGBOIQOKIRKTDCHP 240
181 ALERLLWRACRGFLIASPRELEQPLEHPVTGEPATWMTFLISYWGBOIQOKIRKTDCHP 240
241 CHVFPFLOEVRRAAGLVLPKGRLLPAPPPDRLRIQEETRLAQELRDVRGNOALRA 300
241 CHVFPFLOEVRRAAGLVLPKGRLLPAPPPDRLRIQEETRLAQELRDVRGNOALRA 300
301 AVYALNOCVSTTHKCLIAEAWCSVRDLPALQELRDSSMBEGSVSAVAHRIPCRDMPPT 360
301 AVYALNOCVSTTHKCLIAEAWCSVRDLPALQELRDSSMBEGSVSAVAHRIPCRDMPPT 360
361 LIRTRFTASFGIIVDRYGVGRYQVNPAPYITITFPFLFAVMPGDVGHGLMFLFALAM 420
361 LIRTRFTASFGIIVDRYGVGRYQVNPAPYITITFPFLFAVMPGDVGHGLMFLFALAM 420
421 VLAENRPVAKAQNIEWOTFFRGRYLLMLLGLFSIYTGFIYNECFSRATSIIPSGMSVAA 480
421 VLAENRPVAKAQNIEWOTFFRGRYLLMLLGLFSIYTGFIYNECFSRATSIIPSGMSVAA 480
481 MANQSGWSDAFIAQHTMLTLDPNVTGVLGPPYFGIDPIWLSLAANHLSPNSFKMKMSVI 540
481 MANQSGWSDAFIAQHTMLTLDPNVTGVLGPPYFGIDPIWLSLAANHLSPNSFKMKMSVI 540
541 LGVHMAFGVVLGVNHNHFGQRHLLLETLPELFLGLFGLVFLVLYKWLVCWAARA 600
541 LGVHMAFGVVLGVNHNHFGQRHLLLETLPELFLGLFGLVFLVLYKWLVCWAARA 600
601 ASPSLIHFIHNNFLSHSPSNRLLYPROQVQATLVVLALAMVPIILLGTPLHLHRRHR 660
601 ASPSLIHFIHNNFLSHSPSNRLLYPROQVQATLVVLALAMVPIILLGTPLHLHRRHR 660
661 RLRRPADRQENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLHQAIHTI 720
661 RLRRPADRQENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLHQAIHTI 720
721 EFCGCVSNTASYLRWALSALAHQSEVLWAMVNRIGLGLRGVGVAAVLPVIFAAPA 780
721 EFCGCVSNTASYLRWALSALAHQSEVLWAMVNRIGLGLRGVGVAAVLPVIFAAPA 780
781 VMTVAILLVMEGLSAPFLHALRLHWVEFQNKFSYGTGYKLSPP 822
781 VMTVAILLVMEGLSAPFLHALRLHWVEFQNKFSYGTGYKLSPP 822

RESULT 2

US-09-618-304B-2
; Sequence 2, Application US/09618304B
; Patent No. 677537
; GENERAL INFORMATION:
; APPLICANT: Stashenko, Philip
; APPLICANT: Li, Yi-Ping
; TITLE OF INVENTION: Osteoclast Proton Pump Subunit
; FILE REFERENCE: 1564.1006-001
; CURRENT APPLICATION NUMBER: US/09/618,304B
; CURRENT FILING DATE: 2006-07-18
; PRIOR APPLICATION NUMBER: 08/605,378
; PRIOR FILING DATE: 1996-02-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Homo sapiens

US-09-618-304B-2
Query Match 100.0%; Score 4266; DB 2; Length 822;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGSFMRSEVALVQLFPLTAAAYTCVSRIGELGLVEFRDLNASVAFQRRFVVVWVRCE 60
DB 1 MGSFMRSEVALVQLFPLTAAAYTCVSRIGELGLVEFRDLNASVAFQRRFVVVWVRCE 60
QY 61 LEKTTFFLOEVRRAAGLVLPKGRLLPAPPPDRLRIQEETRLAQELRDVRGNOALRA 120
DB 61 LEKTTFFLOEVRRAAGLVLPKGRLLPAPPPDRLRIQEETRLAQELRDVRGNOALRA 120
QY 121 QHQQLHAAVLRQGHPEPQAAAAHTDASERTPLQAQGGPHQDLRVNFVAGAVPHKAP 180
DB 121 QHQQLHAAVLRQGHPEPQAAAAHTDASERTPLQAQGGPHQDLRVNFVAGAVPHKAP 180
QY 181 ALERLLWRACRGFLIASPRELEQPLEHPVTGEPATWMTFLISYWGBOIQOKIRKTDCHP 240
DB 181 ALERLLWRACRGFLIASPRELEQPLEHPVTGEPATWMTFLISYWGBOIQOKIRKTDCHP 240
QY 241 CHVFPFLOEVRRAAGLVLPKGRLLPAPPPDRLRIQEETRLAQELRDVRGNOALRA 300
DB 241 CHVFPFLOEVRRAAGLVLPKGRLLPAPPPDRLRIQEETRLAQELRDVRGNOALRA 300
QY 301 AVYALNOCVSTTHKCLIAEAWCSVRDLPALQELRDSSMBEGSVSAVAHRIPCRDMPPT 360
DB 301 AVYALNOCVSTTHKCLIAEAWCSVRDLPALQELRDSSMBEGSVSAVAHRIPCRDMPPT 360
QY 361 LIRTRFTASFGIIVDRYGVGRYQVNPAPYITITFPFLFAVMPGDVGHGLMFLFALAM 420
DB 361 LIRTRFTASFGIIVDRYGVGRYQVNPAPYITITFPFLFAVMPGDVGHGLMFLFALAM 420
QY 421 VLAENRPVAKAQNIEWOTFFRGRYLLMLLGLFSIYTGFIYNECFSRATSIIPSGMSVAA 480
DB 421 VLAENRPVAKAQNIEWOTFFRGRYLLMLLGLFSIYTGFIYNECFSRATSIIPSGMSVAA 480
QY 481 MANQSGWSDAFIAQHTMLTLDPNVTGVLGPPYFGIDPIWLSLAANHLSPNSFKMKMSVI 540
DB 481 MANQSGWSDAFIAQHTMLTLDPNVTGVLGPPYFGIDPIWLSLAANHLSPNSFKMKMSVI 540
QY 541 LGVHMAFGVVLGVNHNHFGQRHLLLETLPELFLGLFGLVFLVLYKWLVCWAARA 600
DB 541 LGVHMAFGVVLGVNHNHFGQRHLLLETLPELFLGLFGLVFLVLYKWLVCWAARA 600
QY 601 ASPSLIHFIHNNFLSHSPSNRLLYPROQVQATLVVLALAMVPIILLGTPLHLHRRHR 660
DB 601 ASPSLIHFIHNNFLSHSPSNRLLYPROQVQATLVVLALAMVPIILLGTPLHLHRRHR 660
QY 661 RLRRPADRQENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLHQAIHTI 720
DB 661 RLRRPADRQENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLHQAIHTI 720
QY 721 EFCGCVSNTASYLRWALSALAHQSEVLWAMVNRIGLGLRGVGVAAVLPVIFAAPA 780
DB 721 EFCGCVSNTASYLRWALSALAHQSEVLWAMVNRIGLGLRGVGVAAVLPVIFAAPA 780
QY 781 VMTVAILLVMEGLSAPFLHALRLHWVEFQNKFSYGTGYKLSPP 822
DB 781 VMTVAILLVMEGLSAPFLHALRLHWVEFQNKFSYGTGYKLSPP 822

RESULT 3
US-09-949-016-8846
; Sequence 8846, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8846
; LENGTH: 849
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8846

Query Match 99.3%; Score 4235.5; DB 2; Length 849;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 820; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 MGSMPSESEVALVQLFLPTAAAYTCVSRGELGLVEFRDLNASVSAPQRRFVVDVVRCEE 60
DB 20 MGSMPSESEVALVQLFLPTAAAYTCVSRGELGLVEFRDLNASVSAPQRRFVVDVVRCEE 79
QY 61 LEKFTFLOEVRAGLVLPPKGRLPAPPDRLLRIQETERLAQELRDVRGNQQAALRA 120
DB 80 LEKFTFLOEVRAGLVLPPKGRLPAPPDRLLRIQETERLAQELRDVRGNQQAALRA 139
QY 121 QLHQLQHAVALRQGHPEPOLAAHTDGASERTPLQAPGGPHQDLRVNFAVAGAVEPHKAP 180
DB 140 QLHQLQHAVALRQGHPEPOLAAHTDGASERTPLQAPGGPHQDLRVNFAVAGAVEPHKAP 199
QY 181 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEOIGKIRKIDTCFH 240
DB 200 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEOIGKIRKIDTCFH 259
QY 241 CHVFPFLOQSEARLQALQLOOQSQSLOEVLGTERFSLQVGRVLQQLPPGQVQVHKMK 300
DB 260 CHVFPFLOQSEARLQALQLOOQSQSLOEVLGTERFSLQVGRVLQQLPPGQVQVHKMK 319
QY 301 AVYALNQC SVSTTHKCLIAEAWCSVRDLPALQEARLDSMEGVSAAHRIPCRMPPT 360
DB 320 AVYALNQC SVSTTHKCLIAEAWCSVRDLPALQEARLDSMEGVSAAHRIPCRMPPT 379
QY 361 LIETNRFTASFOGIVDYGVRQYQEVNPAPYITITPPFLFAVMFGDVGHGLMFLFALAM 420
DB 380 LIETNRFTASFOGIVDYGVRQYQEVNPAPYITITPPFLFAVMFGDVGHGLMFLFALAM 439
QY 421 VLAENRPAKAAQNEIWTQFFRGRYLLMLGLFSIYTGFIYNECFSRATSIIPSGWSVAA 480
DB 440 VLAENRPAKAAQNEIWTQFFRGRYLLMLGLFSIYTGFIYNECFSRATSIIPSGWSVAA 499
QY 481 MANQSGWSDAFLAQHTMTLDPNVTVGFLGYPFGIDPIWLSLAANHLFLNSFKMKMSVI 540
DB 500 MANQSGWSDAFLAQHTMTLDPNVTVGFLGYPFGIDPIWLSLAANHLFLNSFKMKMSVI 559
QY 541 LGVHMAFGVVLGVFNHVFHQHRLLETLPELTLLGLFGYLVFLVIYKMLCVWAARA 600
DB 560 LGVHMAFGVVLGVFNHVFHQHRLLETLPELTLLGLFGYLVFLVIYKMLCVWAARA 619
QY 601 AS-PSLIIHFNNFLFSHPNSRLLYPROEVQATLVIALAMVPIILLGTPLHLHRRH 659

DB 800 AVMTVAILLVMEGLSAPFLHALRLHWBFQNKFSYGTGYKLSPF 842
RESULT 4
US-09-949-016-8847
; Sequence 8847, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8847
; LENGTH: 849
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8847

Query Match 99.3%; Score 4235.5; DB 2; Length 849;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 820; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 1 MGSMPSESEVALVQLFLPTAAAYTCVSRGELGLVEFRDLNASVSAPQRRFVVDVVRCEE 60
DB 20 MGSMPSESEVALVQLFLPTAAAYTCVSRGELGLVEFRDLNASVSAPQRRFVVDVVRCEE 79
QY 61 LEKFTFLOEVRAGLVLPPKGRLPAPPDRLLRIQETERLAQELRDVRGNQQAALRA 120
DB 80 LEKFTFLOEVRAGLVLPPKGRLPAPPDRLLRIQETERLAQELRDVRGNQQAALRA 139
QY 121 QLHQLQHAVALRQGHPEPOLAAHTDGASERTPLQAPGGPHQDLRVNFAVAGAVEPHKAP 180
DB 140 QLHQLQHAVALRQGHPEPOLAAHTDGASERTPLQAPGGPHQDLRVNFAVAGAVEPHKAP 199
QY 181 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEOIGKIRKIDTCFH 240
DB 200 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEOIGKIRKIDTCFH 259
QY 241 CHVFPFLOQSEARLQALQLOOQSQSLOEVLGTERFSLQVGRVLQQLPPGQVQVHKMK 300
DB 260 CHVFPFLOQSEARLQALQLOOQSQSLOEVLGTERFSLQVGRVLQQLPPGQVQVHKMK 319
QY 301 AVYALNQC SVSTTHKCLIAEAWCSVRDLPALQEARLDSMEGVSAAHRIPCRMPPT 360
DB 320 AVYALNQC SVSTTHKCLIAEAWCSVRDLPALQEARLDSMEGVSAAHRIPCRMPPT 379
QY 361 LIETNRFTASFOGIVDYGVRQYQEVNPAPYITITPPFLFAVMFGDVGHGLMFLFALAM 420
DB 380 LIETNRFTASFOGIVDYGVRQYQEVNPAPYITITPPFLFAVMFGDVGHGLMFLFALAM 439
QY 421 VLAENRPAKAAQNEIWTQFFRGRYLLMLGLFSIYTGFIYNECFSRATSIIPSGWSVAA 480
DB 440 VLAENRPAKAAQNEIWTQFFRGRYLLMLGLFSIYTGFIYNECFSRATSIIPSGWSVAA 499
QY 481 MANQSGWSDAFLAQHTMTLDPNVTVGFLGYPFGIDPIWLSLAANHLFLNSFKMKMSVI 540
DB 500 MANQSGWSDAFLAQHTMTLDPNVTVGFLGYPFGIDPIWLSLAANHLFLNSFKMKMSVI 559
QY 541 LGVHMAFGVVLGVFNHVFHQHRLLETLPELTLLGLFGYLVFLVIYKMLCVWAARA 600
DB 560 LGVHMAFGVVLGVFNHVFHQHRLLETLPELTLLGLFGYLVFLVIYKMLCVWAARA 619
QY 601 AS-PSLIIHFNNFLFSHPNSRLLYPROEVQATLVIALAMVPIILLGTPLHLHRRH 659

Db 620 ASAPGILIHFINMFLSHSPSNRLLYPRQEVVQATLVVLALAMVPIILGLTGPLHLLHRRH 679
Qy 660 RRLRRPADROENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLMHQAIHT 719
Db 680 RRLRRPADROENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLMHQAIHT 739
Qy 720 IEFCGCVSNTASYLRWLWALSIAHAQSLSEVLWAMVMRIGLGLGREGVAAVVLVPIPAF 779
Db 740 IEFCGCVSNTASYLRWLWALSIAHAQSLSEVLWAMVMRIGLGLGREGVAAVVLVPIPAF 799
Qy 780 AVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPP 822
Db 800 AVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPP 842

RESULT 5
US-09-949-016-8275
; Sequence 8275, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8275
; LENGTH: 853
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8275

Query Match 99.3%; Score 4235.5; DB 2; Length 853;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 820; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 1 MGSMFRSEVALVOLFLPTAAAYTCVSRGLGELVFEFRDLNASVSFAFORFVVDVRRCEE 60
Db 24 MGSMFRSEVALVOLFLPTAAAYTCVSRGLGELVFEFRDLNASVSFAFORFVVDVRRCEE 83
Qy 61 LEKTFPFOEVRVRRAGLVLPKGLPAPPDRDLRIQEEERLAQELRDVGRNQOALRA 120
Db 84 LEKTFPFOEVRVRRAGLVLPKGLPAPPDRDLRIQEEERLAQELRDVGRNQOALRA 143
Qy 121 QLHLQLHAAVLRQGHPEQLAAATDGCASERTPLQAPGGPHQDLRVNFVAGAVEPHKAP 180
Db 144 QLHLQLHAAVLRQGHPEQLAAATDGCASERTPLQAPGGPHQDLRVNFVAGAVEPHKAP 203
Qy 181 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATMTFLISYWGEOIGQIRKKTDCFH 240
Db 204 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATMTFLISYWGEOIGQIRKKTDCFH 263
Qy 241 CHVPFPLOQEARLGALQOQOSELQEVLETGERFLSQVLGRVLQLLPGQOVQVHOK 300
Db 264 CHVPFPLOQEARLGALQOQOSELQEVLETGERFLSQVLGRVLQLLPGQOVQVHOK 323
Qy 301 AVYALNQCSTVTHKCLIAEAWCSVRDLPALQEARLDRSSMBEGSVSAVAHRIPCRDMPPT 360
Db 324 AVYALNQCSTVTHKCLIAEAWCSVRDLPALQEARLDRSSMBEGSVSAVAHRIPCRDMPPT 383
Qy 361 LIRNRTFASFGIVDRGVGRYQVSNPAPYTIITFPFLFAMFGDVGHLLMFLPALAM 420
Db 384 LIRNRTFASFGIVDAVGVGRYQVSNPAPYTIITFPFLFAMFGDVGHLLMFLPALAM 443

Qy 421 VLAENRPAVKAQNEIMOTFRGRVYLLLMGLFSYITGTFIYNECFSRATSIIPSGWSVAA 480
Db 444 VLAENRPAVKAQNEIMOTFRGRVYLLLMGLFSYITGTFIYNECFSRATSIIPSGWSVAA 503
Qy 481 MANQSGWSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWSLAANHLSPLNSFKMKMSVI 540
Db 504 MANQSGWSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWSLAANHLSPLNSFKMKMSVI 563
Qy 541 LGVHMAFGVVLGVNHHVHFGQRHLLLETPELTFLGLFGYLVFLVIYKWLVCVWAARA 600
Db 564 LGVHMAFGVVLGVNHHVHFGQRHLLLETPELTFLGLFGYLVFLVIYKWLVCVWAARA 623
Qy 601 AS-PSILIHFINMFLSHSPSNRLLYPRQEVVQATLVVLALAMVPIILGLTGPLHLLHRRH 659
Db 624 ASAPGILIHFINMFLSHSPSNRLLYPRQEVVQATLVVLALAMVPIILGLTGPLHLLHRRH 683
Qy 660 RRLRRPADROENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLMHQAIHT 719
Db 684 RRLRRPADROENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLMHQAIHT 743
Qy 720 IEFCGCVSNTASYLRWLWALSIAHAQSLSEVLWAMVMRIGLGLGREGVAAVVLVPIPAF 779
Db 744 IEFCGCVSNTASYLRWLWALSIAHAQSLSEVLWAMVMRIGLGLGREGVAAVVLVPIPAF 803
Qy 780 AVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPP 822
Db 804 AVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPP 846

RESULT 6
US-09-949-016-8276
; Sequence 8276, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8276
; LENGTH: 853
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8276

Query Match 99.3%; Score 4235.5; DB 2; Length 853;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 820; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 1 MGSMFRSEVALVOLFLPTAAAYTCVSRGLGELVFEFRDLNASVSFAFORFVVDVRRCEE 60
Db 24 MGSMFRSEVALVOLFLPTAAAYTCVSRGLGELVFEFRDLNASVSFAFORFVVDVRRCEE 83
Qy 61 LEKTFPFOEVRVRRAGLVLPKGLPAPPDRDLRIQEEERLAQELRDVGRNQOALRA 120
Db 84 LEKTFPFOEVRVRRAGLVLPKGLPAPPDRDLRIQEEERLAQELRDVGRNQOALRA 143
Qy 121 QLHLQLHAAVLRQGHPEQLAAATDGCASERTPLQAPGGPHQDLRVNFVAGAVEPHKAP 180
Db 144 QLHLQLHAAVLRQGHPEQLAAATDGCASERTPLQAPGGPHQDLRVNFVAGAVEPHKAP 203
Qy 181 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATMTFLISYWGEOIGQIRKKTDCFH 240
Db 204 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATMTFLISYWGEOIGQIRKKTDCFH 263

ORGANISM: Human
US-09-949-016-7202

Query Match 79.6%; Score 3394; DB 2; Length 750;
Best Local Similarity 90.7%; Pred. No. 0;
Matches 664; Conservative 6; Mismatches 18; Indels 44; Gaps 3;

```
QY 131 VLROGHEPQAAHTDGAERTPLLOAGGP-----161
D 16 LQARNFQSALALGG---SPLVLPRPAPFPGCGQCGQALDPCPCWAGSCPGGLVG 71
QY 162 -----HODLRNVFVAGVEPHKAPALERLLWRACRGFLIASFRELEOPLEHPVTG 211
D 72 EAGQEVGAALPSFVAASVAGAVEPHKAPALERLLWRACRGFLIASFRELEOPLEHPVTG 131
QY 212 EPATWMTFLISYWGEOIGQIRKIDTCFCHVFPFLOQBEARLGAQQOQSOELQEV 271
D 132 EPATWMTFLISYWGEOIGQIRKIDTCFCHVFPFLOQBEARLGAQQOQSOELQEV 191
QY 272 GETERFLSOVLGRVQLQLPPGOVQVHKMAVYALNOCVSTTHKCLIAEAMCSVRDLPA 331
D 192 GETERFLSOVLGRVQLQLPPGOVQVHKMAVYALNOCVSTTHKCLIAEAMCSVRDLPA 251
QY 332 LQEARLDSMREGVSAVAHRIPCRDMPTLIRNRTASFOGIVDRYGVGRYQEVNPAPY 391
D 252 LQEARLDSMREGVSAVAHRIPCRDMPTLIRNRTASFOGIVDRYGVGRYQEVNPAPY 311
QY 392 TITTEPFLFVAVMGDVGHLMLPALAMVLAENRPAKAAQNEIWQTFRGRYLLLMG 451
D 312 TITTEPFLFVAVMGDVGHLMLPALAMVLAENRPAKAAQNEIWQTFRGRYLLLMG 371
QY 452 LFSIYTGFIYNECFSRATSIFFSGSWAAMANOSGSDAFQAQHTMTLDPNVTGVFLGP 511
D 372 LFSIYTGFIYNECFSRATSIFFSGSWAAMANOSGSDAFQAQHTMTLDPNVTGVFLGP 431
QY 512 YPGIDPISLAANHLFLNSFKMKSIVLGIVVHMAFGVVLGVFNHVFQGRHLLLETL 571
D 432 YPGIDPISLAANHLFLNSFKMKSIVLGIVVHMAFGVVLGVFNHVFQGRHLLLETL 491
QY 572 PELTLLGLFGLVFLVIYKWCVAARAAS-PSILIHFINFLFSDPSNRLLYPRQEV 630
D 492 PELTLLGLFGLVFLVIYKWCVAARAAS-PSILIHFINFLFSDPSNRLLYPRQEV 551
QY 631 VQATLVVLAAMPVILLGTPLHLHRRRRRRPADROENKAGLLDLPASVNGWSS 690
D 552 VQATLVVLAAMPVILLGTPLHLHRRRRRRPADROENKAGLLDLPASVNGWSS 611
QY 691 DEKAGGLDDEBEAEIVSEVLMHQAHTIEFCLGCVSNTASYLRWLWALSLAQAQSEVL 750
D 612 DEKAGGLDDEBEAEIVSEVLMHQAHTIEFCLGCVSNTASYLRWLWALSLAQAQSEVL 671
QY 751 WAMVMRIIGLGRGVGAAVLVPIFAAPAVMTVAILLVMEGLSAFLHALRLHWVFQNK 810
D 672 WAMVMRIIGLGRGVGAAVLVPIFAAPAVMTVAILLVMEGLSAFLHALRLHWVFQNK 731
QY 811 FYSGTGYKLSPP 822
D 732 FYSGTGYKLSPP 743
```

RESULT 9

5196526-1

PATENT NO. 5196526

APPLICANT: BEAMAN, KENNETH D.

TITLE OF INVENTION: CDNA CLONE FOR T-CELL SUPPRESSOR INDUCER

FACTOR

NUMBER OF SEQUENCES: 3

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/442,116

FILING DATE: 28-NOV-1989

SEQ ID NO.1:

LENGTH: 855

5196526-1

Query Match 50.2%; Score 2143.5; DB 7; Length 855;
Best Local Similarity 51.4%; Pred. No. 3.8e-208;
Matches 441; Conservative 129; Mismatches 233; Indels 55; Gaps 15;

```
QY 1 MGSFMRSEVALVQLFLPTAAAYTCVSRIGELGLVEFRDLNASVSFAQRFFVVDVWRCCE 60
D 1 MGSFMRSEVALVQLFLPTAAAYTCVSRIGELGLVEFRDLNASVSFAQRFFVVDVWRCCE 60
QY 61 LKTTFTTFOEVRVRLGLVLPKPKRLPAPPPDLRLRIBETRLRLAQLDELVRGNQOALRA 120
D 61 LKTTFTTFOEVRVRLGLVLPKPKRLPAPPPDLRLRIBETRLRLAQLDELVRGNQOALRA 120
QY 121 QL-----HQQLAAAVLRQCHE-----PQLAAHTDGAERTPLLOAGGPQHDLRVN 168
D 121 NLELVEYTHMLRVTKTFLKRNVEFPEYEEPALENDSLDYS-CMORLGA-----KLG 174
QY 169 FVAGAVEPHKAPALERLLWRACRGFLIASFRELEOPLEHPVTGEPATWMTFLISYWGEOI 228
D 175 FVAGAVEPHKAPALERLLWRACRGFLIASFRELEOPLEHPVTGEPATWMTFLISYWGEOI 234
QY 229 GQIKIRKIDTCFCHVFPFLOQBEARLGAQQOQSOELQEVGETERFLSOVLGRVQL 288
D 235 GQIKIRKIDTCFCHVFPFLOQBEARLGAQQOQSOELQEVGETERFLSOVLGRVQL 294
QY 289 LPPGOVQVHKMAVYALNOCVSTTHKCLIAEAMCSVRDLPAQEARLDSMREG--VS 346
D 295 LPPGOVQVHKMAVYALNOCVSTTHKCLIAEAMCSVRDLPAQEARLDSMREG--VS 354
QY 347 AVAHRIPCRDMPTLIRNRTASFOGIVDRYGVGRYQEVNPAPYTIITFPFLFAVMFGD 406
D 355 AVAHRIPCRDMPTLIRNRTASFOGIVDRYGVGRYQEVNPAPYTIITFPFLFAVMFGD 414
QY 407 VGHGILLMFLPALAMVLAENRPAKAAQNEIWQTFRGRYLLLMGLFSIYTGFIYNECF 466
D 415 VGHGILLMFLPALAMVLAENRPAKAAQNEIWQTFRGRYLLLMGLFSIYTGFIYNECF 473
QY 467 RATSIFFPSGWSVAAMANOSG-----WSDAFLAQHTMTLDPNVTGVFLGPYPCGI 516
D 474 RATSIFFPSGWSVAAMANOSG-----WSDAFLAQHTMTLDPNVTGVFLGPYPCGI 533
QY 517 DPTSLAANHLFLNSFKMKSIVLGIVVHMAFGVVLGVFNHVFQGRHLLLETLPELTF 576
D 534 DPTSLAANHLFLNSFKMKSIVLGIVVHMAFGVVLGVFNHVFQGRHLLLETLPELTF 593
QY 577 LLGLFGLVFLVIYKWCVAARA-ASPSILIHFINFLFSDPSNRLLYPRQEVVQATL 635
D 594 LLGLFGLVFLVIYKWCVAARA-ASPSILIHFINFLFSDPSNRLLYPRQEVVQATL 652
QY 636 VVLAAMPVILLGTPLHLHRRRR-----LRRRPADROENKAGLLDLPAS 684
D 653 VVLAAMPVILLGTPLHLHRRRR-----LRRRPADROENKAGLLDLPAS 708
QY 685 VNGWSSDEKAGGLDDEBEAEIVSEVLMHQAHTIEFCLGCVSNTASYLRWLWALSLAHA 744
D 709 VNGWSSDEKAGGLDDEBEAEIVSEVLMHQAHTIEFCLGCVSNTASYLRWLWALSLAHA 764
QY 745 QLSVLMWAMVRIGLGRGVGAAVLVPIFAAPAVMTVAILLVMEGLSAFLHALRLHW 804
D 765 QLSVLMWAMVRIGLGRGVGAAVLVPIFAAPAVMTVAILLVMEGLSAFLHALRLHW 821
QY 805 VEFQNKFGYKLSPP 822
D 822 VEFQNKFGYKLSPP 839
```

RESULT 10

US-09-618-304B-3

Sequence 3, Application US/09618304B

Patent No. 677537

GENERAL INFORMATION:

APPLICANT: Stashenko, Philip

APPLICANT: Li, Yi-Ping

```

; TITLE OF INVENTION: Ostecoclast Proton Pump Subunit
; FILE REFERENCE: 1564.1006-001
; CURRENT APPLICATION NUMBER: US/09/618,304B
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 08/605,378
; PRIOR FILING DATE: 1996-02-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence
; NAME/KEY: VARIANT
; LOCATION: (1)...(847)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-618-304B-3

Query Match      42.2%; Score 1801.5; DB 2; Length 847;
Best Local Similarity 46.8%; Pred. No. 2.1e-173;
Matches 391; Conservative 45; Mismatches 384; Indels 15; Gaps 6;

Qy   1 MGSMFRSEVALVQLFLPHTAAATCVSRIGELGLVFPRDLNANVSAPQRFFVDVWRCEE 60
Db   1 MGXLFSEEXLXLQLFLXXXAAYKCVSLXELGVAFRDLNXXVFXQRFVKVEVRCEE 60

Qy   61 LKFTFPLQEVRRAGLVLPKPGRLPAPPDRLLRIQEETRLAQELRDVRGNQAALRA 120
Db   61 MDXKLFPXXKEIRAXXXXXXXXXXXXXXPDMXXXXXXXXXXEKKELKEIXNXQALXX 120

Qy   121 QLHLQLHAAVLRLQGHEPQLAAAAHTDGASERTPLL-QAPCGPHQDILRVNFVAGAVEPKHA 179
Db   121 XFXXLXXXXXILRXXXXXXXXXXXDXDXEXXLLXXXXXXGXXXXXLRXXFVAGXIXXXXX 180

Qy   180 PALERLLWACRGFLIASFRELEQPLeHPVTGPATMTFLISYMGEGIQKIRKITDCF 239
Db   181 PXFERMLWRXCRCGXGXXXXXXEXEPLEXPTGDXXXXXXXFXIXFGDXQXXXXXXIKEXF 240

Qy   240 HCVFPLQOEEARLGALQLOOQSOELQVLETGERFLSVLGRVLQLLPGQOVVHKM 299
Db   241 XXXXPFXGXXXXXXGXXXXXXGXXXXXXDQVLXTTXXXXXXVXXXXXXXXXXXXXIXVMX 300

Qy   300 KAVYLALNQCSVSTTHKCLIAEAWCVRDLPALQEAALRDSMEEG--VSAVAHRIPCRDM 357
Db   301 KAIVYXNLNCXIXYTKCLIAEWXCVXDLYXXQXALRXXXXXXGXXVXXIYXRXXXXXX 360

Qy   358 PPTLIPTNRFTASFGQIDRYGVGYOEVNPAPYTIITPPFLFAVMFGDVGHGLLMFLFA 417
Db   361 PPTXXXTNXFTXXFOXIVD-YGIXYXEINPAPYTIITPPFLFAVMFGDXGHGLMXLFLFA 419

Qy   418 LAMVLAENRAPVAAQNEIWOTFFRGYRLLLLMGLFSIYTGFIYNECFSTRATSIFPSGWS 477
Db   420 XXMVLEEXRXXXXXXNEXFXXXFXGYYXLLMGXFSIYTGFIYNDYCFSSXXXIFXSXWS 479

Qy   478 VAAVANQSOWSDAFLAQHTWLTLDPNVTVGVGLGYPFGIDPIWLSLAANHLSFLNSFKMKM 537
Db   480 VXXMXXXXXXEXEXXXXXXLXLPXVGXVFGYFGIDPIWXXAXNKLXFLNSFKMKM 539

Qy   538 SVILGVVHMAGVYLVGVNHFHQGRHLLLETLPBELTFLGLFGYLVPLVIYKWLVCMA 597
Db   540 SVILGIITHKFGVKLXXFNHXXFXXXXXXPEXFMXXLFGYLVKLIYKXXXXXX 599

Qy   598 A--RAASPSTLIHFIMNPLFSHSPS-NRLLYPQEVQVQATLVIALAMVPILLGTPLHL 654
Db   600 AXXXXXXXPSXLIHFIMNPLFSXXXXXNMXYXXQXXIQXXLVVXALXXVPXMLXXPLXL 659

Qy   655 LHRHR-----RLRRRPADQRQENKAGLLDLPDASNNGSWSSDEEKAGGLDDEAEAL 706
Db   660 XXXXXXXXXXXXXXXXXXXXFXXXXXXXXXXXXXXXXXXXXXXXXXXEDXXXDF 719

Qy   707 VPSEVLMMHQAIHTIEFCGLGSNTASYRLRWALSALAHQOLSEVLWAMVNRIGLGLGREVG 766

```

Db 286 ----- 285
QY 523 AANHLSFLNSFKMKHVSILGVVHMAFGVVLGVFNHVFQGRHLLLETLPELTFLLGLPG 582
Db 286 ----- 285
QY 583 YLVFLVIYKWLVCMAARAASPSILIHFINMFLSHS-PSNRLLYPRQEVVQATLVVLALA 641
Db 286 ----- 285
QY 642 MYPIILLGTP--LHLHRRHRRRLRRPADROENKAGLLDLPDASVNGSSDEEKAGGLD 699
Db 323 SVPMMLLIKPFILASHRSKSQLQASRIQEDATENIEGSSSP-SRSRGORTSADTHGALD 381
QY 700 DEEBEELVPSEVLHQAIHTIEFCGCVSNTASYLRLWALSALAHQAOLSEVLWAMVRIGL 759
Db 382 DHGE-EFNFQGVDFVQALHTIIEYCGCISNTASYLRLWALSALAHQAOLSEVLWAMVRIGL 440
QY 760 ---GLGREYGAUVLVPIFAAFVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTG 816
Db 441 QTRGWGGIVGVFI---IFAVFAVLITVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTG 495
QY 817 YKLSPP 822
Db 496 YKFSPP 501

RESULT 12

US-09-270-767-43227
; Sequence 43227, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 1999-03-17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43227
; LENGTH: 515
; TYPE: PRT
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-43227

Query Match 20.9%; Score 892.5; DB 2; Length 515;
Best Local Similarity 40.0%; Pred. No. 2.1e-81;
Matches 191; Conservative 81; Mismatches 162; Indels 43; Gaps 7;
QY 23 YTCVRLGELGLVFEFRDLNASVSQRRFVVDVWRCEELEKTFTLQBEVVRAGLVLP 82
Db 43 YTSVSELGETGVQVRDLNVNVAQRKEVTEVRCDLELRKIRVIEIKDGIIVLPI 102
QY 83 KGRLP-APPDRLLRIQ---EETE---RLAQ-----ELRDVRGNQALRAQ 121
Db 103 QDDIPRAPNPREIIDLEAHLKTESMIELAQNEVMKSNYLELTELKRVLENTQGFSD 162
QY 122 LHQQLHAALVLRQHEPQLAAHTDGAERTLLQAPGPHQDLAVNVAGAVEPHKAPA 181
Db 163 QEVLLDSS-NRAGGDNDAQAHRG-----RLGFVAGVINRERVA 202
QY 182 LERLLWRACRGLIASFRELEOPLHPVTGEPATWNTFLISYWGQIGQIKRIKIDCFHC 241
Db 203 FERMLWRISRGVFLKRSDLDEPLNDPATGHPYIKTVFVAFQGEQLKKNRIKVCCTGFA 262
QY 242 HVFPFLOQBEARLQALQLOQSQBELQVGLTERFLSQLGRVLQQLPPGQVQVHKMA 301
Db 263 SLYPSPSHNREERENVRTRLEDLKLVSQTEHRSRVLATVSKNLPSPSMVWKMA 322
QY 302 VYLAHQCSVSTHKLITAEACWSVRDLPALQEARLDSMEEG--VSAVAHRIPCDMPP 359
Db 323 IYHTLNLNFMNDVTKKCLIGECVWPTNDLPVQKALSDGSAAGVSTIPSLNVIDTNEQPP 382

QY 360 TLIRNRTASTQGIQVDRVGVGYOVNPPAPVTIITTEPFLPAMFGDVGHLLMELFALA 419
Db 363 TFRNKNKTRGPQNLIDAYGVASYRECPALYTCITFFFLPAMFGDGLGHGUILVDFGAW 442
QY 420 MYLAENRPAVKAQAONEIMQTFPRGRYLLMLGLFSIYTGFIYNECFSRATSPPSGW 476
Db 443 MYLCERKLARINGGEIWNIPFGGRYIILLMGLFAMYTGLVYNDVFSKSMNLFGRW 499

RESULT 13

US-09-248-796A-17451
; Sequence 17451, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17451
; LENGTH: 386
; TYPE: PRT
; ORGANISM: *Candida albicans*
US-09-248-796A-17451

Query Match 17.8%; Score 757.5; DB 2; Length 386;
Best Local Similarity 41.0%; Pred. No. 7e-68;
Matches 161; Conservative 71; Mismatches 132; Indels 29; Gaps 7;
QY 433 QNEIMQTFPRGRYLLMLGLFSIYTGFIYNECFSRATSPPSGWSVAAMANQSGSDAFL 492
Db 5 KDEIFDMAYTGRYVLLLMGVFSMYTGFYNDVFSISMSIFKSGWE-----WPEKEN 55

QY 493 AQHTMLTLDPNVTGVFLGPYPGIDPIWLSLAANHLSPFNFSFKMKSVILGVVHMAFGVVL 552
Db 56 VGET-----IYAKYVGTYSIGLDPAMHGTENALLFSNSYKMKLSILMGYHMSYVVF 108
QY 553 GVFNHVFQGRHRLLETLPELTFLLGLFVGVLFVIVYKWLVCMAARAASPSILH-FTN 611
Db 109 SLVNTYFNSMIDVIGNFIPGLLFQGIQFGLSLCIVTKNSVDWVFATGRQPEGLNMLIN 168
QY 612 MFLFSHPSPN--RLLYPRQEVVQATLVVLALAMVPIILLGTPLHLHLHRRRLRRRPPAD 669
Db 169 MFL---QPGDVPEPLYSQSTIQVFLLLIALICVPMLLLVKPLYM-----KQLEKEANQ 220
QY 670 QEENKAGLLDLPDASVNGSSDEEKAGGLDDEEBELVPSEVLHQAIHTIEFCGCVSN 729
Db 221 HGSYSQALNDEESGAVQGEQENAAEDDDDEHEHNFQDITMHIQVHTIEFCLNCVSH 280
QY 730 TASYLRLWALSALAHQAOLSEVLWAMVRIGLGREGVGAUVLVPIFAAFVMTVAIVL 789
Db 281 TASYLRLWALSALAHQAOLSTVLWSMT--IGNAFGPTGLIGTFMVVFLFAMVFLSVCLVV 338
QY 790 MEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPP 822
Db 339 MEGTSAMLSHLRLHWVESMSKYFEGGSAGFATF 371

RESULT 14

US-09-270-767-44518
; Sequence 44518, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44518
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-44518

Query Match 16.5%; Score 704; DB 2; Length 380;
Best Local Similarity 39.4%; Pred. No. 1.9e-62;
Matches 150; Conservative 70; Mismatches 147; Indels 14; Gaps 4;
QY 88 APPRDLRLTQETERLAQRLDVRGNQQAALRAQLHQLHAAVLRGHEPQLAAHTDG 147
DB 4 APOPREMIDLEATFEKLENELEVNQNAEALKRNFLELTTELKHLRKTQVFFDEMAQNQN 63
QY 148 ASERTPLLAQPG---GPHODLRVNFVAGAVEPHKAPALERLLWRACRGFLIASFRELEQ 203
DB 64 EDEQAQLLGEGRASQPGQNLKLGFGVAGVILRERLPAFERMLWRACRGVFLRQAMET 123
QY 204 PLEHPVTGEPATWMTFLISYWGEGIQKIRKIDTDFCHVFPF----LQOEERLGLAQ 259
DB 124 PLEDPNTGDQVHKSVPIIFQGDQLKTRVKKICGFRATLYPCPEAPADRRWAMGVMT 183
QY 260 LQOQSBLQELVETERFLSQLGRVLQLLPPGQVQVHKMAVYLALNQCSVSTTHKCLI 319
DB 184 I-----EDLNTVLGQTQDHRHRLVAAAANKLNWFKVKIKAIYHTLNLFLNLDVTQKCLI 239
QY 320 AEAWCSVRDLPALQEARLDSSMEEG--VSAVAHRIPCRDMPPTLIRNRETASFOGIVDR 377
DB 240 AECVPLLDIETQLAURRTERSGSVPPILNRMTFENPPYTNRTNKTAKPAQALIDA 299
QY 378 YGVRQYQEVNPAPYTIITFPFLFAVMEGDVGHGLMLFALAMVLAENRPAVKAQNEIW 437
DB 300 YGVASYRENNPAPYTIITFPFLFAVMEGDVGHGLMALFGLWMIRKEKGLAAQKTNEIW 359
QY 438 QTFPRGRYLLMLGLFSIYTG 458
DB 360 NIFFGGRYIIFLMGVFSMYTG 380

RESULT 15
US-09-270-767-33142
; Sequence 33142, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33142
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-33142

Query Match 12.0%; Score 513; DB 2; Length 257;
Best Local Similarity 43.8%; Pred. No. 2.5e-43;
Matches 113; Conservative 45; Mismatches 68; Indels 32; Gaps 9;
QY 584 LVFLVYKVLVCVWAARAA-----SPSLIHFINMFLFSHP-----SNRLLYPRQEVV 631
DB 1 MVFMFEKWS-YNARTSFQPTPGCAPSVLIIFINMLFKNTPEPPKGCNEFWFESQPL 59
QY 632 QATLVVALAMVPILLGTPLHLHRRRLRRPADRQBEENK-AGLLDLFDASV---NG 687
DB 60 QKAPVLIALCCIPWMLLGKPLYI-----KPTRKNKAHANHGQLTGNIELAEGETPLPTG 114
QY 688 WSSDEEKAGGLDDEEABLVSEVLHMQAIHTIEFCLGCVSNTASYLRLWALSIAHAQLS 747

DB 115 FSGNEENAGGAHGHDDDEPM--SEIYIHOAIHTTIEYVLSTISHTASYLRLWALSIAHAQLS 172
QY 748 EVLWAMVMRIGL---GLGREVGVAADVVLVPIFAAFVMTVAILLVMEGLSAFLHALRLHW 804
DB 173 EVLWQMVLSLGLKMSGVGGAIG-----LPIIFGAWCLFTLAILVMEGLSAFLHTLRLHW 227
QY 805 VEFONKPYSGTGYKLSGF 822
DB 228 VEFMSKPYEGMGYAFQPF 245

Search completed: June 29, 2006, 13:09:59
Job time : 56 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2006, 13:09:14 ; Search time 181 Seconds
(without alignments)
2103.661 Million cell updates/sec

Title: US-10-783-519-2
Perfect score: 4266
Sequence: 1 MGSFSESEVALVQLFLPTA.....HWVEFQNFYSGTGYKLSDF 822

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:

- 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|----------------------|
| 1 | 4266 | 100.0 | 822 | 5 | US-10-783-519-2 |
| 2 | 3131.5 | 73.4 | 614 | 4 | US-10-145-012-2 |
| 3 | 3131.5 | 73.4 | 614 | 6 | US-11-126-866-2 |
| 4 | 3131.5 | 73.4 | 614 | 6 | US-11-126-841A-2 |
| 5 | 3127.5 | 73.3 | 614 | 4 | US-10-145-012-13 |
| 6 | 3127.5 | 73.3 | 614 | 5 | US-10-489-725-5 |
| 7 | 3127.5 | 73.3 | 614 | 6 | US-11-126-866-13 |
| 8 | 3127.5 | 73.3 | 614 | 6 | US-11-126-841A-13 |
| 9 | 2917.5 | 68.4 | 643 | 4 | US-10-264-049-2656 |
| 10 | 2066 | 48.4 | 818 | 4 | US-10-495-446-32 |
| 11 | 1918.5 | 45.0 | 841 | 5 | US-10-874-706-20 |
| 12 | 1801.5 | 42.2 | 847 | 5 | US-10-783-519-3 |
| 13 | 1720 | 40.3 | 855 | 5 | US-10-714-995-14 |
| 14 | 1720 | 40.3 | 855 | 6 | US-11-097-143-2052 |
| 15 | 1701 | 39.9 | 834 | 6 | US-11-097-143-4199 |
| 16 | 1701 | 39.9 | 834 | 6 | US-11-097-143-41700 |
| 17 | 1539.5 | 36.1 | 703 | 4 | US-10-495-446-28 |
| 18 | 1539.5 | 36.1 | 935 | 4 | US-10-369-493-5810 |
| 19 | 1538 | 36.1 | 814 | 6 | US-11-097-143-36642 |
| 20 | 1501 | 35.2 | 1030 | 4 | US-10-369-493-5811 |
| 21 | 1497 | 35.1 | 844 | 6 | US-11-097-143-41496 |
| 22 | 1491.5 | 35.0 | 865 | 4 | US-10-369-493-5572 |
| 23 | 1491.5 | 35.0 | 865 | 4 | US-10-369-493-5573 |
| 24 | 1406.5 | 33.0 | 873 | 4 | US-10-369-493-6070 |
| 25 | 1375 | 32.2 | 1236 | 4 | US-10-369-493-6203 |
| 26 | 1325.5 | 31.1 | 820 | 4 | US-10-437-963-196349 |
| 27 | 1318 | 30.9 | 840 | 4 | US-10-369-493-22409 |

28 1293.5 30.3 784 4 US-10-437-963-196347
29 1285.5 30.1 558 5 US-10-450-763-39739
30 1280 30.0 794 4 US-10-495-446-31
31 1268.5 29.7 783 4 US-10-437-963-152242
32 1217 28.5 957 4 US-10-369-493-3848
33 1208 28.3 877 5 US-10-485-986-22
34 1144 26.8 890 4 US-10-369-493-1874
35 1108.5 26.0 821 4 US-10-017-161-2144
36 1108.5 26.0 821 4 US-10-292-798-1790
37 1102 25.8 702 4 US-10-389-566-1961
38 1097.5 25.7 513 4 US-10-094-749-2211
39 1084 25.4 792 4 US-10-437-963-172172
40 1010 23.7 535 3 US-09-949-029-93
41 1010 23.7 535 3 US-09-949-029-94
42 964 22.6 589 4 US-10-369-493-12518
43 673.5 15.8 372 4 US-10-424-599-255903
44 670 15.7 459 4 US-10-424-599-187383
45 630 14.8 340 4 US-10-424-599-258221

ALIGNMENTS

RESULT 1

US-10-783-519-2
; Sequence 2, Application US/10783519
; Publication No. US2005006448A1
; GENERAL INFORMATION:
; APPLICANT: Stashenko, Philip
; APPLICANT: Li, Yi-Ping
; TITLE OF INVENTION: Osteoclast Proton Pump Subunit
; FILE REFERENCE: 1564.1006-001
; CURRENT APPLICATION NUMBER: US/10783,519
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/09/618,304
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 08/605,378
; PRIOR FILING DATE: 1996-02-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-783-519-2

Query Match 100.0%; Score 4266; DB 5; Length 822;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSFSESEVALVQLFLPTAAAYTCVSRGELGLVFRDLNASVSFAQRFVVDVWRCEE 60
DB 1 MGSFSESEVALVQLFLPTAAAYTCVSRGELGLVFRDLNASVSFAQRFVVDVWRCEE 60

QY 61 LEKTFTFLOEVRAGLVLPKGRPLPAPPDRLLRIQETTERLAQLRDLVRGNQQAALRA 120
DB 61 LEKTFTFLOEVRAGLVLPKGRPLPAPPDRLLRIQETTERLAQLRDLVRGNQQAALRA 120

QY 121 QLHQLHAAVLRLQGHPEPOLAAHTDGASERTPLLOAGPGPHODLRVNFVAGAVEPHKAP 180
DB 121 QLHQLHAAVLRLQGHPEPOLAAHTDGASERTPLLOAGPGPHODLRVNFVAGAVEPHKAP 180

QY 181 ALERLLWRCRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEGIQKIRKTDTCFH 240
DB 181 ALERLLWRCRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEGIQKIRKTDTCFH 240

QY 241 CHVFPFLOEVRAGLVLPKGRPLPAPPDRLLRIQETTERLAQLRDLVRGNQQAALRA 300
DB 241 CHVFPFLOEVRAGLVLPKGRPLPAPPDRLLRIQETTERLAQLRDLVRGNQQAALRA 300

QY 301 AVYLAALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLDSMREGVSAVAHRIPCRDMPT 360
DB 301 AVYLAALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLDSMREGVSAVAHRIPCRDMPT 360

QY 361 LIRTNRTASFOGIVDRYGVGRYQEVNPAPYTIITPPFLFAVMFGDVGHGLLMFLPALAM 420
DB |||||
DB 361 LIRTNRTASFOGIVDRYGVGRYQEVNPAPYTIITPPFLFAVMFGDVGHGLLMFLPALAM 420
QY 421 VLAENRPAKAAQNEIWTQFFRGRYLLLLMGLFSIYTGFIYNECFSRATSIIPSGMSVAA 480
DB |||||
DB 421 VLAENRPAKAAQNEIWTQFFRGRYLLLLMGLFSIYTGFIYNECFSRATSIIPSGMSVAA 480
QY 481 MANQSGWSDAFLAQTMTLTDPNVTGVLGPPFGDIPWISLAANHLSPNSFKMKMSVI 540
DB |||||
DB 481 MANQSGWSDAFLAQTMTLTDPNVTGVLGPPFGDIPWISLAANHLSPNSFKMKMSVI 540
QY 541 LGVWMAFGVILGVFNHVFQGRHRLLETLBELTFLGLGVLVFLVIYKWLVCWAARA 600
DB |||||
DB 541 LGVWMAFGVILGVFNHVFQGRHRLLETLBELTFLGLGVLVFLVIYKWLVCWAARA 600
QY 601 ASPSILIHFINMFLFSHSPSNRLLYPRQEVVQATLVVLAAMVPILLGTPLHLLHRHR 660
DB |||||
DB 601 ASPSILIHFINMFLFSHSPSNRLLYPRQEVVQATLVVLAAMVPILLGTPLHLLHRHR 660
QY 661 RLRRRPAQOEKAGLLDLPASVNGWSSDEKAGGLDDEAEALVPSVLVMPHQAHTI 720
DB |||||
DB 661 RLRRRPAQOEKAGLLDLPASVNGWSSDEKAGGLDDEAEALVPSVLVMPHQAHTI 720
QY 721 EFCLGCVSNTASYRLMALSLAHAQLSEVLWAMVRIGLGLGREVGVAAVLVLPFAFA 780
DB |||||
DB 721 EFCLGCVSNTASYRLMALSLAHAQLSEVLWAMVRIGLGLGREVGVAAVLVLPFAFA 780
QY 781 VMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYSGTGKLSPPF 822
DB |||||
DB 781 VMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYSGTGKLSPPF 822

RESULT 2

US-10-145-012-2
; Sequence 2, Application US/10145012
; Publication No. US20030124614A1
; GENERAL INFORMATION:
; APPLICANT: UTKU et al.
; TITLE OF INVENTION: T-CELL MEMBRANE PROTEIN (TIRC7), PEPTIDES AND ANTIBODIES DERIVED
; FILE REFERENCE: 4400-0105P
; CURRENT APPLICATION NUMBER: US/10/145,012
; CURRENT FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-145-012-2

Query Match 73.4%; Score 3131.5; DB 4; Length 614;
Best Local Similarity 99.7%; Pred. No. 2.4e-265;
Matches 605; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 217 MTFLISYWGEGIQGKIRKITDCFHCHVPPFLQOEERLALGALQLOOQSOELQEVGETER 276
DB |||||
DB 1 MTFLISYWGEGIQGKIRKITDCFHCHVPPFLQOEERLALGALQLOOQSOELQEVGETER 60
QY 277 FLSQVLGRVLQLLPPGQVQVHKMAVYLALNQCSTVTHKCLIAEAWCSVRDLPALQEA 336
DB |||||
DB 61 FLSQVLGRVLQLLPPGQVQVHKMAVYLALNQCSTVTHKCLIAEAWCSVRDLPALQEA 120
QY 337 RDSSMEEGVSAVAHRIPCRDMPPTLIRNRTFTASFGIVDRYGVGRYQEVNPAPYTIITF 396
DB |||||
DB 121 RDSSMEEGVSAVAHRIPCRDMPPTLIRNRTFTASFGIVDRYGVGRYQEVNPAPYTIITF 180
QY 397 PFLFAVMFGDVGHGLLMFLPALAMVLAENRPAKAAQNEIWTQFFRGRYLLLLMGLFSIY 456
DB |||||
DB 181 PFLFAVMFGDVGHGLLMFLPALAMVLAENRPAKAAQNEIWTQFFRGRYLLLLMGLFSIY 240
QY 457 TGFYNECFSRATSIIPSGMSVAAANOSGSDAFLAQTMTLTDPNVTGVLGPPFGI 516
DB |||||
DB 241 TGFYNECFSRATSIIPSGMSVAAANOSGSDAFLAQTMTLTDPNVTGVLGPPFGI 300

QY 457 TGFYNECFSRATSIIPSGMSVAAANOSGSDAFLAQTMTLTDPNVTGVLGPPFGI 516
DB |||||
DB 241 TGFYNECFSRATSIIPSGMSVAAANOSGSDAFLAQTMTLTDPNVTGVLGPPFGI 300
QY 517 DPWISLAANHLSPNSFKMKMSVILGVVHMAFGVVLGVFNHVFQGRHRLLETLBELT 576
DB |||||
DB 301 DPWISLAANHLSPNSFKMKMSVILGVVHMAFGVVLGVFNHVFQGRHRLLETLBELT 360
QY 577 LGLFGYLVFLVIYKWLVCWAARAAS-PSILIHFINMFLFSHSPSNRLLYPRQEVVQATL 635
DB |||||
DB 361 LGLFGYLVFLVIYKWLVCWAARAASPSILIHFINMFLFSHSPSNRLLYPRQEVVQATL 420
QY 636 VVLAAMVPILLGTPLHLLHRRLRRRPAQOEKAGLLDLPASVNGWSSDEEKA 695
DB |||||
DB 421 VVLAAMVPILLGTPLHLLHRRLRRRPAQOEKAGLLDLPASVNGWSSDEEKA 480
QY 696 GGLDDEEAEALVPSVLVMPHQAHTIETFCVNTASYRLMALSLAHAQLSEVLWAMVM 755
DB |||||
DB 481 GGLDDEEAEALVPSVLVMPHQAHTIETFCVNTASYRLMALSLAHAQLSEVLWAMVM 540
QY 756 RIGLGLGREVGVAAVLVLPFAFAVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSY 815
DB |||||
DB 541 RIGLGLGREVGVAAVLVLPFAFAVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSY 600
QY 816 GYKLSPPF 822
DB |||||
DB 601 GYKLSPPF 607

RESULT 3

US-11-126-866-2
; Sequence 2, Application US/11126866
; Publication No. US20050220789A1
; GENERAL INFORMATION:
; APPLICANT: UTKU, et al., NALAN
; TITLE OF INVENTION: METHODS OF TREATMENT USING AN ANTIBODY THAT SPECIFICALLY BINDS TO
; FILE REFERENCE: 1472/71099-ZA/JPW/AG
; CURRENT APPLICATION NUMBER: US/11/126,866
; CURRENT FILING DATE: 2005-05-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 614
; TYPE: PRT
; ORGANISM: human
US-11-126-866-2

Query Match 73.4%; Score 3131.5; DB 6; Length 614;
Best Local Similarity 99.7%; Pred. No. 2.4e-265;
Matches 605; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 217 MTFLISYWGEGIQGKIRKITDCFHCHVPPFLQOEERLALGALQLOOQSOELQEVGETER 276
DB |||||
DB 1 MTFLISYWGEGIQGKIRKITDCFHCHVPPFLQOEERLALGALQLOOQSOELQEVGETER 60
QY 277 FLSQVLGRVLQLLPPGQVQVHKMAVYLALNQCSTVTHKCLIAEAWCSVRDLPALQEA 336
DB |||||
DB 61 FLSQVLGRVLQLLPPGQVQVHKMAVYLALNQCSTVTHKCLIAEAWCSVRDLPALQEA 120
QY 337 RDSSMEEGVSAVAHRIPCRDMPPTLIRNRTFTASFGIVDRYGVGRYQEVNPAPYTIITF 396
DB |||||
DB 121 RDSSMEEGVSAVAHRIPCRDMPPTLIRNRTFTASFGIVDRYGVGRYQEVNPAPYTIITF 180
QY 397 PFLFAVMFGDVGHGLLMFLPALAMVLAENRPAKAAQNEIWTQFFRGRYLLLLMGLFSIY 456
DB |||||
DB 181 PFLFAVMFGDVGHGLLMFLPALAMVLAENRPAKAAQNEIWTQFFRGRYLLLLMGLFSIY 240
QY 457 TGFYNECFSRATSIIPSGMSVAAANOSGSDAFLAQTMTLTDPNVTGVLGPPFGI 516
DB |||||
DB 241 TGFYNECFSRATSIIPSGMSVAAANOSGSDAFLAQTMTLTDPNVTGVLGPPFGI 300
QY 517 DPWISLAANHLSPNSFKMKMSVILGVVHMAFGVVLGVFNHVFQGRHRLLETLBELT 576

```
Db 301 DPTWLAANHLSPNSFKMMSVILGVVHAFGVVLGVFNHVFQGRHRLLETLPELTFF 360
QY 577 LLGLFGYLVFLVIYKWLVCWAARAAS-PSILIHFINMFLFSHSPSNRLLYPRQEVQATL 635
Db 361 LLGLFGYLVFLVIYKWLVCWAARAASAPSILIHFINMFLFSHSPSNRLLYPRQEVQATL 420
QY 636 VVLALAMVPILLGLTGPLHLHRRRRLRRRPPADQENKAGLLDLPDASVNGWSSDDEKA 695
Db 421 VVLALAMVPILLGLTGPLHLHRRRRLRRRPPADQENKAGLLDLPDASVNGWSSDDEKA 480
QY 696 GGLDDEEAEALVPSEVLVMPHQAIIHTIEFCLGCVSNNTASYLRILWALSIAHAQLSEVLWAMVM 755
Db 481 GGLDDEEAEALVPSEVLVMPHQAIIHTIEFCLGCVSNNTASYLRILWALSIAHAQLSEVLWAMVM 540
QY 756 RIGLGLRGVGAHVLPVPIFAAFVAVMTVAILLVMEGLSAFLHALRLHWVEFQNKFYSGT 815
Db 541 RIGLGLRGVGAHVLPVPIFAAFVAVMTVAILLVMEGLSAFLHALRLHWVEFQNKFYSGT 600
QY 816 GYKLSPP 822
Db 601 GYKLSPP 607
```

RESULT 4

US-11-126-841A-2

; Sequence 2, Application US/11126841A

; Publication No. US20050271659A1

; GENERAL INFORMATION:

; APPLICANT: UTKU, et al., NALAN

; TITLE OF INVENTION: METHODS OF TREATMENT USING AN ANTIBODY THAT SPECIFICALLY BINDS TO

; FILE REFERENCE: 1472/71099-ZB/JPW/AG

; CURRENT APPLICATION NUMBER: US/11/126,841A

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 2

; LENGTH: 614

; TYPE: PRT

; ORGANISM: human

US-11-126-841A-2

```
Query Match 73.4%; Score 3131.5; DB 6; Length 614;
Best Local Similarity 99.7%; Pred. No. 2.4e-265;
Matches 605; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
```

```
QY 217 MTELISYWGEOIGQIKIRKIDTCFCHVFPFLOQEEARLQALQOQSQBELQEVLGSTER 276
Db 1 MTELISYWGEOIGQIKIRKIDTCFCHVFPFLOQEEARLQALQOQSQBELQEVLGSTER 60
QY 277 FLSQVLGRVLQQLPPGQOVQVHKMAVYALNQCSTVTHKCLIAEAWCSVRDLPALQEAL 336
Db 61 FLSQVLGRVLQQLPPGQOVQVHKMAVYALNQCSTVTHKCLIAEAWCSVRDLPALQEAL 120
QY 337 RDSMEEGVSVAHRIPCRDMPPTLIRTNRTFASFGQIVDYGVRGYQEVNPAPYTIITF 396
Db 121 RDSMEEGVSVAHRIPCRDMPPTLIRTNRTFASFGQIVDYGVRGYQEVNPAPYTIITF 180
QY 397 PFLFAMFGDVGHGLLMFLPALAMVLAENRPAVKAQAQNEIWQTFGRGYLLLLMGLFSIY 456
Db 181 PFLFAMFGDVGHGLLMFLPALAMVLAENRPAVKAQAQNEIWQTFGRGYLLLLMGLFSIY 240
QY 457 TGFYINCEFSRATSIFFPSGWSVAAMANQSGWSDAFLAQHTMTLTDNPVTCGVFLGYPYFPGI 516
Db 241 TGFYINCEFSRATSIFFPSGWSVAAMANQSGWSDAFLAQHTMTLTDNPVTCGVFLGYPYFPGI 300
QY 517 DPTWLAANHLSPNSFKMMSVILGVVHAFGVVLGVFNHVFQGRHRLLETLPELTFF 576
Db 301 DPTWLAANHLSPNSFKMMSVILGVVHAFGVVLGVFNHVFQGRHRLLETLPELTFF 360
QY 577 LLGLFGYLVFLVIYKWLVCWAARAAS-PSILIHFINMFLFSHSPSNRLLYPRQEVQATL 635
Db 361 LLGLFGYLVFLVIYKWLVCWAARAASAPSILIHFINMFLFSHSPSNRLLYPRQEVQATL 420
QY 577 LLGLFGYLVFLVIYKWLVCWAARAAS-PSILIHFINMFLFSHSPSNRLLYPRQEVQATL 635
```

```
Db 361 LLGLFGYLVFLVIYKWLVCWAARAASAPSILIHFINMFLFSHSPSNRLLYPRQEVQATL 420
QY 636 VVLALAMVPILLGLTGPLHLHRRRRLRRRPPADQENKAGLLDLPDASVNGWSSDDEKA 695
Db 421 VVLALAMVPILLGLTGPLHLHRRRRLRRRPPADQENKAGLLDLPDASVNGWSSDDEKA 480
QY 696 GGLDDEEAEALVPSEVLVMPHQAIIHTIEFCLGCVSNNTASYLRILWALSIAHAQLSEVLWAMVM 755
Db 481 GGLDDEEAEALVPSEVLVMPHQAIIHTIEFCLGCVSNNTASYLRILWALSIAHAQLSEVLWAMVM 540
QY 756 RIGLGLRGVGAHVLPVPIFAAFVAVMTVAILLVMEGLSAFLHALRLHWVEFQNKFYSGT 815
Db 541 RIGLGLRGVGAHVLPVPIFAAFVAVMTVAILLVMEGLSAFLHALRLHWVEFQNKFYSGT 600
QY 816 GYKLSPP 822
Db 601 GYKLSPP 607
```

RESULT 5

US-10-145-012-13

; Sequence 13, Application US/10145012

; Publication No. US20030124614A1

; GENERAL INFORMATION:

; APPLICANT: UTKU et al.

; TITLE OF INVENTION: T-CELL MEMBRANE PROTEIN (TIRC7), PEPTIDES AND ANTIBODIES DERIVE

; FILE REFERENCE: 4400-0105P

; CURRENT APPLICATION NUMBER: US/10/145,012

; CURRENT FILING DATE: 2002-05-13

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 13

; LENGTH: 614

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-145-012-13

```
Query Match 73.3%; Score 3127.5; DB 4; Length 614;
Best Local Similarity 99.5%; Pred. No. 5.3e-265;
Matches 604; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
```

```
QY 217 MTELISYWGEOIGQIKIRKIDTCFCHVFPFLOQEEARLQALQOQSQBELQEVLGSTER 276
Db 1 MTELISYWGEOIGQIKIRKIDTCFCHVFPFLOQEEARLQALQOQSQBELQEVLGSTER 60
QY 277 FLSQVLGRVLQQLPPGQOVQVHKMAVYALNQCSTVTHKCLIAEAWCSVRDLPALQEAL 336
Db 61 FLSQVLGRVLQQLPPGQOVQVHKMAVYALNQCSTVTHKCLIAEAWCSVRDLPALQEAL 120
QY 337 RDSMEEGVSVAHRIPCRDMPPTLIRTNRTFASFGQIVDYGVRGYQEVNPAPYTIITF 396
Db 121 QDSMEEGVSVAHRIPCRDMPPTLIRTNRTFASFGQIVDYGVRGYQEVNPAPYTIITF 180
QY 397 PFLFAMFGDVGHGLLMFLPALAMVLAENRPAVKAQAQNEIWQTFGRGYLLLLMGLFSIY 456
Db 181 PFLFAMFGDVGHGLLMFLPALAMVLAENRPAVKAQAQNEIWQTFGRGYLLLLMGLFSIY 240
QY 457 TGFYINCEFSRATSIFFPSGWSVAAMANQSGWSDAFLAQHTMTLTDNPVTCGVFLGYPYFPGI 516
Db 241 TGFYINCEFSRATSIFFPSGWSVAAMANQSGWSDAFLAQHTMTLTDNPVTCGVFLGYPYFPGI 300
QY 517 DPTWLAANHLSPNSFKMMSVILGVVHAFGVVLGVFNHVFQGRHRLLETLPELTFF 576
Db 301 DPTWLAANHLSPNSFKMMSVILGVVHAFGVVLGVFNHVFQGRHRLLETLPELTFF 360
QY 577 LLGLFGYLVFLVIYKWLVCWAARAAS-PSILIHFINMFLFSHSPSNRLLYPRQEVQATL 635
Db 361 LLGLFGYLVFLVIYKWLVCWAARAASAPSILIHFINMFLFSHSPSNRLLYPRQEVQATL 420
QY 636 VVLALAMVPILLGLTGPLHLHRRRRLRRRPPADQENKAGLLDLPDASVNGWSSDDEKA 695
Db 421 VVLALAMVPILLGLTGPLHLHRRRRLRRRPPADQENKAGLLDLPDASVNGWSSDDEKA 480
```

| | | | |
|----|-----|---|-----|
| Qy | 696 | GGLDDEERAEVLVPSEVLMHQAIHTTIEFCLGCVNSTASYLRRLWALSIAHAQLSEVLWAMVM | 755 |
| | | | |
| Dd | 481 | GGLDDEERAEVLVPSEVLMHQAIHTTIEFCLGCVNSTASYLRRLWALSIAHAQLSEVLWAMVM | 540 |
| | | | |
| Qy | 756 | RIGLGLRGVEGVAAVVLVPIPAAPAVMTVAILLVMGLSAPFLHALRHLHWVEFQNKFYSCT | 815 |
| | | | |
| Dd | 541 | RIGLGLRGVEGVAAVVLVPIPAAPAVMTVAILLVMGLSAPFLHALRHLHWVEFQNKFYSCT | 600 |
| | | | |
| Qy | 816 | GKLSPPF | 822 |
| | | | |
| Dd | 601 | GKLSPPF | 607 |

```

RESULT 6
US-10-489-725-5
; Sequence 5, Application US/10489725
; Publication No. US20050048067A1
; GENERAL INFORMATION:
; APPLICANT: Utku, Nalan
; TITLE OF INVENTION: Peptides capable of modulating immune response
; FILE REFERENCE: GE19A46/P-PWO
; CURRENT APPLICATION NUMBER: US/10/489,725
; CURRENT FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: 60/322,896
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/322,895
; PRIOR FILING DATE: 2001-09-17
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-489-725-5

```

| Query Match | 73.3% | Score 3127.5 | DB 5 | Length 614 |
|-----------------------|-------|--|----------|------------|
| Best Local Similarity | 99.5% | Pred. No. 5.3e-265 | | |
| Matches | 604 | Conservative 1 | Indels 1 | Gaps 1 |
| QY | 217 | MTFLISYWGEOIGOKIRKITDCEHCHVPFFLQOEERALGNLQOLQOQSOEQLQEVLTGETTER | 276 | |
| DB | 1 | MTFLISYWGEOIGOKIRKITDCEHCHVPFFLQOEERALGNLQOLQOQSOEQLQEVLTGETTER | 60 | |
| QY | 277 | FLSQVLGRVLQLLPPGQVQVHKMAVYLLALNQCSTVTHKCLIAEAWCSVRDLPALQEAL | 336 | |
| DB | 61 | FLSQVLGRVLQLLPPGQVQVHKMAVYLLALNQCSTVTHKCLIAEAWCSVRDLPALQEAL | 120 | |
| QY | 337 | RDSMEGVSVAHRIPCRDMPPTLIRTNRFPTASFGQIVDRYGVGRYQEVNPAPYTIITF | 396 | |
| DB | 121 | QDSMEGVSVAHRIPCRDMPPTLIRTNRFPTASFGQIVDAVGUGRYQEVNPAPYTIITF | 180 | |
| QY | 397 | PFLFAVMFGDVGHGLLMPFLPALAMVLAENRPAVKAQAQNEIWQTFPRGRYLLLLMGLFSYI | 456 | |
| DB | 181 | PFLFAVMFGDVGHGLLMPFLPALAMVLAENRPAVKAQAQNEIWQTFPRGRYLLLLMGLFSYI | 240 | |
| QY | 457 | TGTYNCECFSRATSIFFSGGWSVAAMANOSGWSDAFLAQHTMLTLDPNTVTGTVFLGYPFGI | 516 | |
| DB | 241 | TGTYNCECFSRATSIFFSGGWSVAAMANOSGWSDAFLAQHTMLTLDPNTVTGTVFLGYPFGI | 300 | |
| QY | 517 | DPWLSLAANHLSFNSFKMKNSVILGVVHMAFGVVLGVFNHVHFCQRHLLLETLPELTF | 576 | |
| DB | 301 | DPWLSLAANHLSFNSFKMKNSVILGVVHMAFGVVLGVFNHVHFCQRHLLLETLPELTF | 360 | |
| QY | 577 | LLGLFGYLVFLVIYKVLVCWMAAAS-PSIIITHFINMFLFSHPSPSNRLLYPRQEVVQATL | 635 | |
| DB | 361 | LLGLFGYLVFLVIYKVLVCWMAAASAPSIILHFINMFLFSHPSPSNRLLYPRQEVVQATL | 420 | |
| QY | 636 | VVLALAMVPILLTGTPHLHLHRHRRRLRRPADROENKAGLLDLPDASVNGWSDEEKA | 695 | |
| DB | 421 | VVLALAMVPILLTGTPHLHLHRHRRRLRRPADROENKAGLLDLPDASVNGWSDEEKA | 480 | |
| QY | 696 | GGLDDBEAEIPLSPSEVLHMQAIIHTIEFCLGCVCVSNTASVYLRIWALSIAHQAQISEVLWAWYM | 755 | |

```

Db      481  GGLDDEBEAEIIVPSEVLVHQAIHTIEFCLGCVSNTASTYLRNLWALSLAHQAQLSEVLWAMVM 540
Qy      756  RIGLGLREGVGVAAVLVPIFAAFVAMTVATLLVMEGLSASFHLARLHLHWVEFQNKFSYSGT 815
Db      541  RIGLGLREGVGVAAVLVPIFAAFVAMTVATLLVMEGLSASFHLARLHLHWVEFQNKFSYSGT 600
Qy      816  GYKLSPPF 822
Db      601  GYKLSPPF 607

RESULT 7
US-11-126-866-13
; Sequence 13, Application US/11126866
; Publication No. US20050220789A1
; GENERAL INFORMATION:
; APPLICANT: UTKU, et al., NALAN
; TITLE OF INVENTION: METHODS OF TREATMENT USING AN ANTIBODY THAT SPECIFICALLY
; TITLE OF INVENTION: AN EXTRACELLULAR DOMAIN OF A TIRC7 MEMBRANE PROTEIN
; FILE REFERENCE: 1472/71099-ZA/JPW/AG
; CURRENT APPLICATION NUMBER: US/11/126.866
; CURRENT FILING DATE: 2005-05-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-126-866-13

```

| Query Match | 73.3% | Score 3127.5 | DB 6 | Length 614 |
|-----------------------|----------------|---|----------|------------|
| Best Local Similarity | 99.5% | Pred. No. 5.3e-265 | | |
| Matches 604 | Conservative 1 | Mismatches 1 | Indels 1 | Gaps 1 |
| QY | 217 | MTFLISYWGSGIQGKIRKIDTCDFCHVFPFLQOEAREALGALQOLQOQSOEQLQEVLTGETER | 276 | |
| Db | 1 | MTFLISYWGSGIQGKIRKIDTCDFCHVFPFLQOEAREALGALQOLQOQSOEQLQEVLTGETER | 60 | |
| QY | 277 | FLSQVLGRVLQLLPPGGQVQVHKMAVTLALNQCSTVTHKCLIAEAMCSVRDLPALQAL | 336 | |
| Db | 61 | FLSQVLGRVLQLLPPGGQVQVHKMAVTLALNQCSTVTHKCLIAEAMCSVRDLPALQAL | 120 | |
| QY | 337 | RDSSMEGVSAVAHRIPCRMPPTLIIRNPTASFGQIVDRYGVGRYOENPAPVTIITF | 396 | |
| Db | 121 | QDSSMEGVSAVAHRIPCRMPPTLIIRNPTASFGQIVDAYGVGRYOENPAPVTIITF | 180 | |
| QY | 397 | PELFVAMFQVGVGHGLMFLFALAMVLAENRPVAKAAQNEIWOTFPRGRYLLMLGLPSIY | 456 | |
| Db | 181 | PELFVAMFQVGVGHGLMFLFALAMVLAENRPVAKAAQNEIWOTFPRGRYLLMLGLPSIY | 240 | |
| QY | 457 | TGFIYNECFSRATSIIPSGSHVAAMANQSGSDAFLAQHTMLTLDPNVTGVFLGYPYPGI | 516 | |
| Db | 241 | TGFIYNECFSRATSIIPSGSHVAAMANQSGSDAFLAQHTMLTLDPNVTGVFLGYPYPGI | 300 | |
| QY | 517 | DPIWSLAANHLSPNSFKMKMSVILGVVHMAFGVVLGVFNHVFQGRHRLLETLPETLF | 576 | |
| Db | 301 | DPIWSLAANHLSPNSFKMKMSVILGVVHMAFGVVLGVFNHVFQGRHRLLETLPETLF | 360 | |
| QY | 577 | LLGLFGYLVFLVIYKWLCVWAARAAS-PSLLIHFINMFLPSHSPSNRLLYPROEVVQATL | 635 | |
| Db | 361 | LLGLFGYLVFLVIYKWLCVWAARAASPSLLIHFINMFLPSHSPSNRLLYPROEVVQATL | 420 | |
| QY | 636 | VWLALAMVPTILLGCTPLHLHRRRRRRRRPADRQENKAGLLDLPDASVNGWSDESKA | 695 | |
| Db | 421 | VWLALAMVPTILLGCTPLHLHRRRRRRRRPADRQENKAGLLDLPDASVNGWSDESKA | 480 | |
| QY | 696 | GGLDDEEAEALVPSEVLHMQAIHTIFBCLGCVSNSTASYLRILWALSLAHQAQSEVLWAMVM | 755 | |
| Db | 481 | GGLDDEEAEALVPSEVLHMQAIHTIFBCLGCVSNSTASYLRILWALSLAHQAQSEVLWAMVM | 540 | |

QY 756 RIGLGLRGVGAHVLPFAAFVAVMTVAILLVMEGLSAPFLHALRLHWVEFQNKFSYGT 815
Db |||||||
QY 541 RIGLGLRGVGAHVLPFAAFVAVMTVAILLVMEGLSAPFLHALRLHWVEFQNKFSYGT 600
Db |||||||

QY 816 GYKLSPP 822
Db |||||||

QY 601 GYKLSPP 607
Db |||||||

RESULT 8

US-11-126-841A-13
; Sequence 13, Application US/11126841A
; Publication No. US20050271659A1
; GENERAL INFORMATION:
; APPLICANT: UTIKU, et al., NALAN
; TITLE OF INVENTION: METHODS OF TREATMENT USING AN ANTIBODY THAT SPECIFICALLY BINDS TO
; FILE OF INVENTION: AN EXTRACELLULAR DOMAIN OF A TIRC7 MEMBRANE PROTEIN
; FILE REFERENCE: 1472/71099-ZB/JPW/AG
; CURRENT APPLICATION NUMBER: US/11/126,841A
; CURRENT FILING DATE: 2005-05-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-126-841A-13

Query Match 73.3%; Score 3127.5; DB 6; Length 614;
Best Local Similarity 99.3%; Pred. No. 5.3e-265;
Matches 604; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 217 MTFLLSYWGEIQGKIRKIDCFCHVFPFLOQEEARLQALQOQSQBELQVIGETER 276
Db |||||||

QY 1 MTFLLSYWGEIQGKIRKIDCFCHVFPFLOQEEARLQALQOQSQBELQVIGETER 60
Db |||||||

QY 277 FLQSQVLRVQLLPPGQVQVHKMAVYLAALNQCSTVTHKCLIAEAWCSVRDLPALQAL 336
Db |||||||

QY 61 FLQSQVLRVQLLPPGQVQVHKMAVYLAALNQCSTVTHKCLIAEAWCSVRDLPALQAL 120
Db |||||||

QY 337 ROSSMEEGSAVAHRIPCRDMPTLIRTNFTASFOGIVDGRYGVQVNPAPYIITF 396
Db |||||||

QY 121 QSSMEEGSAVAHRIPCRDMPTLIRTNFTASFOGIVDGRYGVQVNPAPYIITF 180
Db |||||||

QY 397 PFLFVAVWFGDVGHGLLMFLFALANVLAENRPVAKAONEIWQFFRGYRLLLLMGLFSY 456
Db |||||||

QY 181 PFLFVAVWFGDVGHGLLMFLFALANVLAENRPVAKAONEIWQFFRGYRLLLLMGLFSY 240
Db |||||||

QY 457 TGFYNECFSRATSIFFSGWSVAAMANQSGWSDAFLAQHTMTLTDPNVTGVFLGPYPFGI 516
Db |||||||

QY 241 TGFYNECFSRATSIFFSGWSVAAMANQSGWSDAFLAQHTMTLTDPNVTGVFLGPYPFGI 300
Db |||||||

QY 517 DPTWSLAANHLSTLNSFKKMSVILGVWMAFGVILGVFNHVFQGRHLLLTLELTF 576
Db |||||||

QY 301 DPTWSLAANHLSTLNSFKKMSVILGVWMAFGVILGVFNHVFQGRHLLLTLELTF 360
Db |||||||

QY 577 LLGLFGYLVFLVYKWLCVWAAAS-PSILIHFINMFLPSHSPSNRLLYPROEVVQATL 635
Db |||||||

QY 361 LLGLFGYLVFLVYKWLCVWAAASPSILIHFINMFLPSHSPSNRLLYPROEVVQATL 420
Db |||||||

QY 636 VVLALAMVPILLGTLPLHLHRRRLRRPADQENKAGLLDLPDASVNGWSSDEEKA 695
Db |||||||

QY 421 VVLALAMVPILLGTLPLHLHRRRLRRPADQENKAGLLDLPDASVNGWSSDEEKA 480
Db |||||||

QY 696 GGLDDEEAELVPSEVLMHQAHTIEFCLGCVNTASYLRLWALSALHAQLSEVLWVM 755
Db |||||||

QY 481 GGLDDEEAELVPSEVLMHQAHTIEFCLGCVNTASYLRLWALSALHAQLSEVLWVM 540
Db |||||||

QY 756 RIGLGLRGVGAHVLPFAAFVAVMTVAILLVMEGLSAPFLHALRLHWVEFQNKFSYGT 815
Db |||||||

QY 541 RIGLGLRGVGAHVLPFAAFVAVMTVAILLVMEGLSAPFLHALRLHWVEFQNKFSYGT 600
Db |||||||

RESULT 9

US-10-264-049-2656
; Sequence 2656, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2656
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (77)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (318)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (598)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-2656

Query Match 68.4%; Score 2917.5; DB 4; Length 643;
Best Local Similarity 99.3%; Pred. No. 1.5e-246;
Matches 559; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 67 FLOEEVRRAGLVLPKGRLPAPPRDLRLRIQETERLAQELDRVGNQALRAQLHLQ 126
Db |||||||

QY 35 FLOEEVRRAGLVLPKGRLPAPPRDLRLRIQETERLAQELDRVGNQALRAQLHLQ 94
Db |||||||

QY 127 LHAVALRQGHQPQLAAAHDTGASERTPLLOAPGGPHQDLRVNFVAGAVEPHKAPALERLL 186
Db |||||||

QY 95 LHAVALRQGHQPQLAAAHDTGASERTPLLOAPGGPHQDLRVNFVAGAVEPHKAPALERLL 154
Db |||||||

QY 187 WRACRGLFIASPRELEQPLEHPVTGPATWMTFLISYWGEOIGQIKRKITDCPHCHVFPF 246
Db |||||||

QY 155 WRACRGLFIASPRELEQPLEHPVTGPATWMTFLISYWGEOIGQIKRKITDCPHCHVFPF 214
Db |||||||

QY 247 LQOEERLALQALQOQSQSQELQEVLTETERFLSQVLRVQLLPPGQVQVHKMAVYLA 306
Db |||||||

QY 215 LQOEERLALQALQOQSQSQELQEVLTETERFLSQVLRVQLLPPGQVQVHKMAVYLA 274
Db |||||||

QY 307 NQCSVSTTHKCLIAEAWCSVRDLPALQALRSSMEEGSAVAHRIPCRDMPTLIRTNR 366
Db |||||||

QY 275 NQCSVSTTHKCLIAEAWCSVRDLPALQALRSSMEEGSAVAHRIPCRDMPTLIRTNR 334
Db |||||||

QY 367 FTASFQGIVDGRYGVQVNPAPYIITFPFLFVAVWFGDVGHGLLMFLPALAMVLAENR 426
Db |||||||

QY 335 FTASFQGIVDGRYGVQVNPAPYIITFPFLFVAVWFGDVGHGLLMFLPALAMVLAENR 394
Db |||||||

QY 427 PAVKAAQNEIWQFFRGYRLLLLMGLFSYITGFIYNECFSRATSIFFSGWSVAAMANQSG 486
Db |||||||

QY 395 PAVKAAQNEIWQFFRGYRLLLLMGLFSYITGFIYNECFSRATSIFFSGWSVAAMANQSG 454
Db |||||||

QY 487 WSDAFLAQHTMTLTDPNVTGVFLGPYPFGIDPISLAANHLSPINSPINSGVILGVVHM 546
Db |||||||

Db 455 WSDAPLAQHTMLTLDENVTVGLPGYPFGIDPIWSLAANHLSFNSFKMKSIVILGVHM 514
QY 547 AGCVVLGVNHHVHFGQRHLLLETLPELTFLGLFGYLVFLVIYKWLVCWAARAAS-PSI 605
Db 515 AFGVVLGVNHHVHFGQRHLLLETLPELTFLGLFGYLVFLVIYKWLVCWAARAASPSI 574
QY 606 LIHFINMFLFSHSPSNRLLYPRQ 628
Db 575 LIHFINMFLFSHSPSNRLLYPRQ 597
RESULT 10
US-10-495-446-32
; Sequence 32, Application US/10495446
; Publication No. US20040214191A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN RAS INTERACTING PROTEIN
; FILE REFERENCE: PB0190
; CURRENT APPLICATION NUMBER: US/10/495,446
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: PCT/US02/35128
; PRIOR FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US 60/332,756
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Aemica Sequence Listing Engine
; SEQ ID NO 32
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-495-446-32

Query Match 48.4%; Score 2066; DB 4; Length 818;
Best Local Similarity 51.1%; Pred. No. 1.le-171;
Matches 426; Conservative 128; Mismatches 225; Indels 54; Gaps 15;
QY 26 VSRIGELGLVPRDLNASVSFAFRFVVDVWRCELEKTFITFLOREVRAGLVLPKGR 85
Db 1 LSALEGKLVQPRDLNQNVSSQRFVGVKRCBELRILVYLQVEITRADIPLEGAS 60
QY 86 LPAPPRDLRIQEBTERLAQRLDVRGNQQAALRAQL-----HQLQHAVALRQGH- 137
Db 61 PPAPPLKHVLEQEQLOKLELRELVTKNKLRLNLELVETHMLRVTKTFLKRVF 120
QY 138 ----POLAAHTDASERTPLQAPGGHQDLRVNFVAGVPHKAPALERLLWRACRGF 193
Db 121 EPTYEFPALENDSLDYS-CHQRLGA-----KLGFVSLIQQGRVEAFERMLWRACKY 174
QY 194 LIASPRELOPLEHPVTGEPATMTFLISYMGEOIGQIKITDCFCFHVFPFLOQEAR 253
Db 175 TLTVAEIDECLEDETEGEVVKWYVFLSFWEQIGHVKKICDCHYCHYPTNTAER 234
QY 254 LGALQLOQOQSOELQVLETERFLSQVLGRVLQLLPPGQVQVHKMAVYALNOCVST 313
Db 235 RBIOGLNTRIQLDVLVTHKTEDYLRVLCKAAESVCSRVQVRKMAIYHMLNWCSPV 294
QY 314 THKCLIAAWCSVRDLPALQALRSSMEEG--VSAVAHRIPCRDMPPTLINTNPTASF 371
Db 295 TNKCLIAEWCVEVDLPGLRRALRESGSGATIPSMNTIPTKETPTLINTNPTSGF 354
QY 372 QGIVDRYGVGRYOENVNPAPTYITTPFLFVNFVGFVGHGLMLFALAMVLAENPAVKA 431
Db 355 QNIVDAYGVGSVREVNPAIFITTPFLFVNFVGFVGHGLMLFALLLVNENPRLSQ 414
QY 432 AQNETWQTFRGRYLLMLGLFSIYTGFIYNECFGRATSIFPSGWSAAMANQSG----- 486
Db 415 SQ-EILRMFFDGRYILLMLGLFSVVTGLIYNDSCFSKSVNLFSGNVSAMYSSSHSPBEQ 473
QY 487 -----WSDAFLAQHTMLTLDENVTVGLPGYPFGIDPIWSLAANHLSFNSFKMKSIVIL 541
Db 474 RQWLWMDSTIRHSRTLQLDNPINPGVPRGYPFGIDPIWNLATNRLTLFNSFKMKSIVIL 533

QY 542 GVVHMAFGVVLGVNHHVHFGQRHLLLETLPELTFLGLFGYLVFLVIYKWLVCWAARA- 600
Db 534 GIFHMTFGVVLGVNHHVHFGQRHLLLETLPELTFLGLFGYLVFLVIYKWLVCWAARA- 593
QY 601 ASPSILIHFINMFLFSHSPSNRLLYPRQEVQATLVVLAALAMVPITLLGLTPLHLLHRRR 660
Db 594 EAPSLIEFINMFLFSHSPSNRLLYPRQEVQATLVVLAALAMVPITLLGLTPLHLLHRRR 652
QY 661 R-----LRRRPADROENKAGLLDLPDASVNGWSSDEEKAGGLDDDEEABLVPS 709
Db 653 RNCFCMSRSRGYTLVRK--DSEEE--VSLIGNODIE--EGNSRME--GCREVTCEEFNF 704
QY 710 EVLMHQAHTITFCIGCVSNATSYLRLLWALSIAHAQLSEVLWAMVIRIGLGLRGVGVAA 769
Db 705 EILMTQALHSIEYICIGCISNTASYLRLLWALSIAHAQLSDVLWAMLMRVGLRVDVTTGV-- 762
QY 770 VVLVPIFAAFVMTVAILLVMEGLSAFLHALRHVVEFQNKFSYSGTGKLSPF 822
Db 763 LLLLPVMAFFAVLTITPILLVMEGLSAFLHAIHLHWVEFQNKFSYSGTGKLSPF 815

RESULT 11

US-10-874-706-20
; Sequence 20, Application US/10874706
; Publication No. US20050048610A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: YANG, Junming
; APPLICANT: YUE, Henry
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURFORD, Neil
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Dying Aina M.
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: HUMAN TRANSPORT PROTEINS
; FILE REFERENCE: PF-0709 PCT
; CURRENT APPLICATION NUMBER: US/10/874,706
; CURRENT FILING DATE: 2004-06-24
; PRIOR APPLICATION NUMBER: US/10/009,328
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/139,923; 60/148,177; 60/149,357; 60/162,287
; PRIOR FILING DATE: 1999-06-17; 1999-08-10; 1999-08-18; 1999-10-28
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 841
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 2762348CD1
US-10-874-706-20

Query Match 45.0%; Score 1918.5; DB 5; Length 841;
Best Local Similarity 49.7%; Pred. No. 1e-158;
Matches 417; Conservative 123; Mismatches 272; Indels 27; Gaps 15;
QY 1 MGSMPRSEVALVQLFLPTAAAYTCVSRIGELGLVEPRDLNASVAFQRRFVVDVWRCE 60
Db 1 MASVRSSEMCLSQLFLQVAAAYCCVAGELGLVQFKDLNNMNVNSFORKFVNEVRCS 60
QY 61 LEKTFITFLOREVRAGLVLPKGRPAPPRDLRIQEBTERLAQRLDVRGNQQAALR- 119
Db 61 LERILRLEDEMONEIVVQLLEKSP-LTPREMITLETVLEKLEGELEQANQQAALQK 119
QY 120 --AQLHQQLHAVALRQGHFQPLAAAHDTGASERTPLQAPGGP-HQDLRVNFVAGVBP 176

Db 120 SFLELTELKYLKKTDQFFETETNADDDFTEDTSGLELLEKAVPAYMTGKLGFTAGCDPT 179
Qy 177 HKAPA-LERLLWRACRGFLIASFRELEPOLEHPVTGEPATMTFLISYWGEOIQKIRKI 235
Db 180 GKMASFERLLMRVCRGNVYLKFSMDAPLEDPTKEEIQKHIIIFYOQEBQLKIKKI 239
Qy 236 TDCFHCHVFPFLOQEEARLGAQLQOQSQELQELVGTETFLSQVLGRVQLQLPPQOVQ 295
Db 240 CDGFRATVYCPPEPAVERREWLESVNVRLLEDITVITQTSRHRQLLQEEAANWHSLIK 299
Qy 296 VHKMKAYVALNQCSTVTHKCLIAEAWCSVRDLPALQEAALRDSMBEGVSAVA--HRI 352
Db 300 VQKMAVYHILLNMCNIDVTQCVIAEWEFPVADATRIKRAL-EQGMELSGSSMAPIMTV 358
Qy 353 PCRDMPPTLIRTRFTASFGQIVDRYGVGRYQEVNPAPIYIITPPFLFAVMFGDVHGLL 412
Db 359 QSKTAPPTNRTNKTAGFQINVDAYGVGSYREINPAPYIITPPFLFAVMFGDCGHGT 418
Qy 413 MFLPALAMVLAENRPAVKAAQNEIWOTFFRGRYLLMLGLFSIYTGFIYNECFSRATSI 472
Db 419 MLLAALMMLNERRLLSOKTDNEIWNTPFHGRYDILLMGIFSIIYTGIIYNDCFSKSLNIF 478
Qy 473 PGSWSVAAMANQSGSDAFQAQHTMTLDPNVTGVFLG-PYPFGIDIPWLSLAANHLSFLN 531
Db 479 GSSWSQPMFRNGTWNTHVMEESIYLQDPAIPGVYFGNYPYFGIDIPWNLASNKLTFLN 538
Qy 532 SFQKMSVILGVHMAFGVVLGVFNHVFQGRHRLLETPELTFLGLFGYLVFLVIYK 591
Db 539 SYRKMSVILGIVQMVFGVILSFNFHIIYFRRTNIIILQFIPMIFILCLFGYLVFMIIFK 598
Qy 592 WLC--VWAARASPSILIHFINFLFSH--PSNRLYPREVVOATLAVLALAMVPTLL 648
Db 599 WCCFDVHVSQA-PSILIHFINFLFNFYSSNAPLYKHQOEVQSFFVVMALISVPMML 657
Qy 649 GTP--LHLHRRRLRRPADRQENKAGLLDLPDASVNGWSDEEKAGGLDDEBAEL 706
Db 658 IKFPILRASHRKSQLQASRIQEDATENIEGDSSTP--SSRSQRTSADTHGALDDHGE-EF 715
Qy 707 VPSEVLHQAIHTIEFCLGCVSNTASYLRWLWALSLAHQALSEVLWAMVMRIGL---GLGR 763
Db 716 NFGDVFHQAIHTIEYCLGICISNTASYLRWLWALSLAHQALSEVLWAMVMRIGLQTRGWG 775
Qy 764 EVCVAAVVLVPIFAAFVMTVAILLVMEGLSAFLHALRLHWVEFONKFYSGTGKLSPP 822
Db 776 IVGVFI-----IFAVFAVLTVAILLIMEGLSAFLHALRLHWVEFONKFYSGTGKLSPP 829

RESULT 12

US-10-783-519-3
; Sequence 3, Application US/10783519
; Publication No. US2005064448A1
; GENERAL INFORMATION:
; APPLICANT: Stashenko, Philip
; TITLE OF INVENTION: Oetoclast Proton Pump Subunit
; FILE REFERENCE: 1564.1006-001
; CURRENT APPLICATION NUMBER: US/10783,519
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/09/618,304
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 08/605,378
; PRIOR FILING DATE: 1996-02-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(847)

; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-783-519-3

Query Match 42.2%; Score 1801.5; DB 5; Length 847;
Best Local Similarity 46.8%; Pred. No. 1.9e-148;
Matches 391; Conservative 45; Mismatches 384; Indels 15; Gaps 6;

Qy 1 MGSMPRESEVALVOLFLPTAAAYTCVSRIGELGLVFRDLNASSVAFQRRPVVDVWRCEE 60
Db 1 MGXLFSEEXLXQLFLXXXAAAYCVSKXLGEXVXFRDLNXXVXXQFQXFXVFEVRRCEE 60
Qy 61 LEXTFTFLOQEVRRAGVLVPPKGRLPAPPRDLRLRQEBTERLAQELRDVRGQQALRA 120
Db 61 MDXILFXFXEIRAXXXXXXXXPPRDMXXXXXXEXXEXXEXXEXXEXXEXXEXXEXXEXX 120
Qy 121 QLHQQLHAAVLRQGHFQPOLAAHTDGASERTPLL-QAPGGPHODLRVNFVAGAVEPHKA 179
Db 121 FXFXLXXXXXILRXXXXXXXAXDXEXXEXXLLXXXXXXGXXKXFXVAGIXXXXXX 180
Qy 180 PALERLLWRACRGFLIASFRELEPOLEHPVTGEPATMTFLISYWGEOIQKIRKIDCF 239
Db 181 PXFERMLWRXCRGXXXXXXEXEXPLEXPVTGDXXXXXXXFXIXFXGQXXXXXXKIXEXF 240
Qy 240 HCVFPPLOQEEARLGAQLQOQSQELQELVGTETFLSQVLGRVQLQLPPQOVHVKM 299
Db 241 XXXXPXXXXXXRXXXXXXXDXLQVLXXTEXXXXXXXVXXXXXXXKXIXVXKM 300
Qy 300 KAVYVALNQCSTVTHKCLIAEAWCSVRDLPALQEAALRDSMBEG--VSAVAHRIPCRDM 357
Db 301 KAIYXXLNCXIXXTXKCLIAEAWCVKDXLXXQXALRXXXXXXGXXVXXIXXXXXXX 360
Qy 358 PPTLIRTRFTASFGQIVDRYGVGRYQEVNPAPIYIITPPFLFAVMFGDVHGLMLFLFA 417
Db 361 PPTXXXTNKTXFXQXIVD-YGIXYXEINPAPYIITPPFLFAVMFGDXGHGKMLFLFA 419
Qy 418 LAMVLAENRPAVKAAQNEIWOTFFRGRYLLMLGLFSIYTGFIYNECFSRATSIFFSGWS 477
Db 420 XXVWLXEXRXXXXXXNEXFXXFXGXXLLMGXFSIYTGIIYNDCFXXXXXIFXSXWS 479
Qy 478 VAAMANQSGSDAFQAQHTMTLDPNVTGVFLGYPYFGIDIPWLSLAANHLSFLNSFKMKM 537
Db 480 VXXMXXXXXXWEXLXXXXXXLXLPXVXGVFXGYPYFGIDIPWXXAXXNLXFLNSFKMKM 539
Qy 538 SVILGVVHMAFGVVLGVFNHVFQGRHRLLETPELTFLGLFGYLVFLVIYKWLVCWA 597
Db 540 SVILGIITHMFGVXLXFXFNHXXFXXXXXXXFXFXFXFXFXFXFXFXFXFXFXFXFX 599
Qy 598 A--RAASPSLIIHFINFLFSHSPS-NRLLYPROEVVOATLAVLALAMVPTLLIGTPLHL 654
Db 600 AXXXXXXPSXLIHFINFLFSXXXXXXNMLYXXQXXIXQXXLVXALXXVPMMLLXXPLXL 659
Qy 655 LHRHRR-----LRLRRPADRQENKAGLLDLPDASVNGWSDEEKAGGLDDEBAEL 706
Db 660 XXXXXRXXXXXXRXXXXXXRXXXXXXRXXXXXXRXXXXXXRXXXXXXRXXXXXXR 719
Qy 707 VPSEVLHQAIHTIEFCLGCVSNTASYLRWLWALSLAHQALSEVLWAMVMRIGLGLRGV 766
Db 720 XXXDXMKHQAIHTIEYCLGICISNTASYLRWLWALSLAHQALSEVLWAMVMRIGLXXXX 779
Qy 767 VAAVVLVPIFAAFVMTVAILLVMEGLSAFLHALRLHWVEFONKFYSGTGKLSPP 821
Db 780 XXXXXFXFXIIPAAFAVXLTVAILLIMEGLSAFLHALRLHWVEFONKFYSGTGKFXP 834

RESULT 13

US-10-714-995-14
; Sequence 14, Application US/10714995
; Publication No. US20060039941A1
; GENERAL INFORMATION:
; APPLICANT: Stam, Lynn
; APPLICANT: Kamdar, Kim
; APPLICANT: Spana, Eric
; APPLICANT: Bachman, Jane

/ TITLE OF INVENTION: Nucleic Acid Sequences from Drosophila melanogaster that encode H
/ TITLE OF INVENTION: Essential for Viability and Uses Thereof

/ FILE REFERENCE: 70201USNP
/ CURRENT APPLICATION NUMBER: US/10/714,995
/ CURRENT FILING DATE: 2003-11-17
/ PRIOR APPLICATION NUMBER: US 60/436,442
/ PRIOR FILING DATE: 2002-12-23
/ NUMBER OF SEQ ID NOS: 63
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 14
/ LENGTH: 855
/ TYPE: PRT
/ ORGANISM: drosophila melanogaster
US-10-714-995-14

Query Match 40.3%; Score 1720; DB 5; Length 855;
Best Local Similarity 43.0%; Pred. No. 2.8e-141;
Matches 374; Conservative 152; Mismatches 267; Indels 76; Gaps 18;
QY 1 MGSFRSERVALVQLPLPTAAAYTCVSRIGELGLVFEFDLNASVSFAFORFVVDVWRCBE 60
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 60
QY 1 MGSFRSEEMALCQLFLQSEAAAYACVSELGELGLVQFRDLNPDVNAFQKFNVEVRCD 60
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 60
QY 61 LEKTFPFLQEEVRRAGLVLPKGRPLAPPDRDLRIQETERLAQELDRVGRNQOALRA 120
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120
QY 61 MERKLYLEKEIKDGPMLDGTGESPEAPQPREMIDLEATPEKLENELEVNQNAEALKR 120
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120
QY 121 QLHQQLQHAVALRQ-----GHEPOLAAAHDTGASERTPLIQ 156
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 156
QY 121 NFLELTTELKHLRKTQVFFDESVPVTVYKSSGYSSSKYRRYPQM--ADQNDEQAQLLG 178
QY 157 APG-----GPHQDLRVNFVAGAVEPHKAPALERLLWRACRGFLIASFRELEQPLEHPVTGE 212
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 212
QY 179 EGVVASQPGQNLKGLFGVAGVILRERLPAFERMLWRACRGVFLRQAMTETPLEDPTNGD 238
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 238
QY 213 PATWMTFLISYWGQIGQIRKITDCFHCVFPF---LQEEARLQALQQLQOQSQELQ 268
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 268
QY 239 QVHKSVPFIFFQDGLKTRVKKICBGRATLYPCPEAPADREMGVMTRI---EDLN 294
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 294
QY 269 EVLGETERFLQVLRVQLPLPPGOVQVHKMAVLAALNQCSTVTHKCLIAEAWCVSD 328
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 328
QY 295 TVLGQTQDHRHVLVAAAKNKNWFVKRKIKAIYHTLNLFLNDVTQKCLIAECVPLLD 354
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 354
QY 329 LPALQEARLDSMBEG--VSAVAHRIPCRDMPPTLIRTNRTFTASFGQIVDRYGVGRYQSV 386
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 386
QY 355 IETIQALARRGTERSGSSVPPILNRMTQFENPTNRTNKNFKYAKQALIDAGVASYREM 414
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 414
QY 387 NPAPYTIITFPPLPAMFGDVGHLLMFLFALAMVLAENRPAVKAQNEIMQTFPRGRYL 446
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 446
QY 415 NPAPYTIITFPPLPAMFGDLGHGAIMALFGLWMIRKEKGLAAQKTQDNEIWNIFPGRYI 474
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 474
QY 447 LLLMGLFSIYTGFIYNECFSRATSIFFPSGWSVAAMANGSWSDAFLAQHTMLTLDPNVTG 506
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 506
QY 475 IFLMGVFSMYTGLIYNDIFPSLNTFGSHWLS-----YKSTVMENKPLQLSPK--G 525
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 525
QY 507 VFLG-PYFPGDIPINSLA-AHLSFLNGFKMKSVLGVVHMAFGVLGVFNHVFHGORH 564
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 564
QY 526 DYEGAPYFPGMDPIQOVAGANKIIFHNAYKMSIIFGVHMFVGVMSWNHTYFRNRI 585
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 585
QY 565 RLLLETLPELTLGLGLGVFLVLYKLCVMA-----ARAASPSILIHFINMFLFS-- 616
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 616
QY 586 SILLYEFIPQLVPLLLFFPMVLLMFIKMKIIPKPAATNDKPYSEACAPSILITFIDMWLFNT 645
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 645
QY 617 -HSPSN--RLLYPRQEVQATLVLAAMVPTLLGLTPLHLRHRRLRRPADRQEN 673
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 673
QY 646 KPPENCCTYPMQGHFIQVLFVLVAVGICPVMLLAKPL-LIMQARKQANVQPI----- 698
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 698
QY 674 KAGLLDLPDASVNGSSDEERAGGDDDEEAEELVPSEVLHQAHTIEFCLGCVSNTASY 733
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 733
QY 699 -AGATS--DAEAGVYNSGSHGGGGHBEEL--SEIFIHQSINTIEYVLGVSHTASY 753
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 753
QY 734 LRLWALSIAHQLSSEVLWAMNRIIGLIGREGVGAVALVPIFAFVMTVAAILVMEGL 793
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 793

Query Match 40.3%; Score 1720; DB 6; Length 855;
Best Local Similarity 43.0%; Pred. No. 2.8e-141;
Matches 374; Conservative 152; Mismatches 267; Indels 76; Gaps 18;
QY 1 MGSFRSERVALVQLPLPTAAAYTCVSRIGELGLVFEFDLNASVSFAFORFVVDVWRCBE 60
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 60
QY 1 MGSFRSEEMALCQLFLQSEAAAYACVSELGELGLVQFRDLNPDVNAFQKFNVEVRCD 60
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 60
QY 61 LEKTFPFLQEEVRRAGLVLPKGRPLAPPDRDLRIQETERLAQELDRVGRNQOALRA 120
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120
QY 61 MERKLYLEKEIKDGPMLDGTGESPEAPQPREMIDLEATPEKLENELEVNQNAEALKR 120
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120
QY 121 QLHQQLQHAVALRQ-----GHEPOLAAAHDTGASERTPLIQ 156
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 156
QY 121 NFLELTTELKHLRKTQVFFDESVPVTVYKSSGYSSSKYRRYPQM--ADQNDEQAQLLG 178
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 178
QY 157 APG-----GPHQDLRVNFVAGAVEPHKAPALERLLWRACRGFLIASFRELEQPLEHPVTGE 212
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 212
QY 179 EGVVASQPGQNLKGLFGVAGVILRERLPAFERMLWRACRGVFLRQAMTETPLEDPTNGD 238
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 238
QY 213 PATWMTFLISYWGQIGQIRKITDCFHCVFPF---LQEEARLQALQQLQOQSQELQ 268
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 268
QY 239 QVHKSVPFIFFQDGLKTRVKKICBGRATLYPCPEAPADREMGVMTRI---EDLN 294
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 294
QY 269 EVLGETERFLQVLRVQLPLPPGOVQVHKMAVLAALNQCSTVTHKCLIAEAWCVSD 328
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 328
QY 295 TVLGQTQDHRHVLVAAAKNKNWFVKRKIKAIYHTLNLFLNDVTQKCLIAECVPLLD 354
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 354
QY 329 LPALQEARLDSMBEG--VSAVAHRIPCRDMPPTLIRTNRTFTASFGQIVDRYGVGRYQSV 386
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 386

```
Db 355 IETIQLALRGTRSSGSPPIILNRMQTFENPPYTNKFTKAFQALIDAYGVASVREM 414
Qy 387 NPAPYTIITPPFLFVAMFGDGHGLMLFLPALAMVLAENRPAPVAKAAQNEIWTQFFRGYRL 446
Db 415 NPAPYTIITPPFLFVAMFGDGHGLMALFGLMWIRKEKGLAAQKTNEIWNIFFGRYI 474
Qy 447 LLMGLFSTYGTGYINECFSRATSIIPSGSVAMANQSGWDAFLAQHTMLTLDNPNVTG 506
Db 475 IFLMGVFSMYTGLIYNDIFSKSLNIFGSHWHL-----YNKSTVMENKFLQLSPK--G 525
Qy 507 VFLG-PYFPGIDPIWLSLA-AHLISFLNSFKMKMSVILGVVHMAFGVVLGVFNHVFQORH 564
Db 526 DYEAGPYFGMDPIWQVAGANKIIFNAYYKKISIIIFGVHMLFGVWMSHNTYFNRI 585
Qy 565 RLLELTPELTPLGLFGYLVFVIVYKWLVCWA-----ARAAASPSILIHFINNLFPS-- 616
Db 586 SLIYEPIQLVFLLLFFVYVLLMFIMFKWIKFAATNDKPYSEACAPSILITFIDMWLFNT 645
Qy 617 -HSPSN--RLIYPRQEVQATLVVIALAMVPILLGCTPLHLHRRRLRRRPPADQREEN 673
Db 646 KPPENCETVMFGQHFQIQLVFLVAVGCCIPVMLLAKPL-LIMQARKQANVOPI----- 698
Qy 674 KAGLLDLPDASVNGSSDEKAGGLDDEEAEIIVPSEVLHMQAIHTIEFCIGCVSNTASY 733
Db 699 -AGATS--DAEAGVNSGSHGGGHEEEEL--SEIFIHQSIHTIEYVLGVSHTASY 753
Qy 734 LRLWALSQAHAQSEVLWAMVMRIGLGLREVGAVALVPIPAFAVMTVAILLVMEGL 793
Db 754 LRLWALSQAHAQAEVLWTMVLISI--GLKQEGPVGGVILTCVPAFWAILTVGILVMEGL 811
Qy 794 SAPHALRLHWVEFQNKFSYGTGYKLSPPF 822
Db 812 SAPHLTRLHLHWVEFQSKFYKQGYAFQPF 840
```

RESULT 15

```
US-11-097-143-41499
; Sequence 41499, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41499
; LENGTH: 834
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-41499
```

```
Query Match 39.9%; Score 1701; DB 6; Length 834;
Best Local Similarity 42.2%; Pred. No. 1.3e-139;
Matches 365; Conservative 147; Mismatches 268; Indels 84; Gaps 18;

Qy 1 MGSMPSESEVALVOLFLPTAAAYTCVSRGELGLVFRDLNANASVASFORFVVDVWRCCE 60
Db 1 MGSMPSESEWALCOMFIQPEAYTSVSELGETCVQVFRDLNVNNAVQRFVFEVRKDE 60
Qy 61 LEKTFTLQOEVRAGLVPKGRLP-APPPRDLRIQ---BETE-----RLAQ----- 106
Db 61 LERKIRVIETIKKDGIVLPDIQDDIPRAPNPREIIDLEAHLEKTESEMIELAQNEVMK 120
Qy 107 -----ELRDVRCNOALRAQLHQLRAVLROGHEPOLAAAHDTGASERTPLQAPG 159
Db 121 SNYLETLRKLENTQGFSDQEVNLNDS--NRAGGNDAAAHQHG----- 166
Qy 160 GPHQDRLVNFVAGABPHKAPALERLLWRACRGFLTASPRELEOPLEHPVTGSPATWMTF 219
Db 167 -----RLGFVAGVINREVRVFAFERMLWIRSGNVFLKRSDDLDEPLNDPATGHIYKTV 220
Qy 220 LISYWGEGIQKIRKTIIDCFCHVFPFLOQEEEARLGALQLOQSOQELQVLETGERFLS 279
Db 221 VAPFQGEQLKNRIKKVCTGTFHASLYPCSSHNEREEEMVRNVRTRLEDKLVLSQTEHRS 280
Qy 280 QVLGRVQLLPPGOVQVHKMAYLALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLDS 339
Db 281 RVLATYSKNLPSWSIMVKMKAIYHTLNLFNMDVTKKCLIGECVPTNDLPVVVKALS 340
Qy 340 SMEEG--VSAAVARIPCROMPPTLIIRNFTASFGQIVDRYGVGRYQEVNPAVPTIITFP 397
Db 341 SAAVGSTIISFLNVIDTNEQPTFNRTKFTGQNLIDAYGVASRECPALYTCITFP 400
Qy 398 FLFVAMFGDVGHGLMLFLPALAMVLAENRPVAKAAQNEIWTQFFRGYRLHLLMGLFSIYT 457
Db 401 FLFVAMFGDLGHGLILVLFCAWVLCERKLARNGGEIWNIFFGCRYIILLMGLFAMYT 460
Qy 458 GFIYNECFSRATSIIPSGHSVAAANQSGWDAFLAQHTMLTLDNPNVTGYFLGPYFGID 517
Db 461 GLVYNDVFSKMNLFGRWF-----NNYNTTTLTNPNLQPPNSSAV--GVYPPGMD 511
Qy 518 PIWSLAANHLISFLNSFKMKMSVILGVVHMAFGVVLGVFNHVFQORHRLLETLPELTFL 577
Db 512 PVWQLADNKIIIFLNSFKMKLSIIFGLVHMFVGVMSVVMFTHEKRVASIFLEFPVQLLFL 571
Qy 578 LGLFGYLVFLVIYKWLVCWMAAA-----SPSILIHFINNLFSSHSP-----SNRLLY 625
Db 572 LLLFGYVFMFMFKWFS--YNARTSPQETPGCAPSVLIMFINMFLFKNTPPKGCNEFMF 630
Qy 626 PRQEVVQATLVVIALAMVILLGCTPLHLHRRRLRRRPADQRENK--AGLLDLPDAS 684
Db 631 ESQPOLQAKAFVLIALLCCIPMMLLGKPLYI-----KFRKNKAHANHNGQLTGNIELAE 685
Qy 685 V---NGWSSDEEKAGGLDDEEAEIIVPSEVLHMQAIHTIEFCIGCVSNTASYLRLWALS 741
Db 686 TPLPTGFGSGNEENAGAGAHGHDDEPM--SEIYIHOAIHTIEYVLSTISHTASYLRLWALS 743
Qy 742 AHAQLSEVLWAMVMRIGL---GLGREVGVAVALVPIFAAFVMTVAILLVMEGLSAFLH 798
Db 744 AHAQLSEVLWQVLSLGLKMSGVGAIG-----LFIIFGAWCLFTLAILVLMGLSAFLH 798
Qy 799 ALRLHWVEFQNKFSYGTGYKLSPPF 822
Db 799 TLRLHWVEFQNKFSYGTGYKLSPPF 822
```

Search completed: June 29, 2006, 13:13:04
Job time : 184 secs

THIS PAGE BLANK (USPTO)

Db 514 GLIKVRDPYFPGVDSWRGSRGELPFLNLSKMKMSILMGVTQWNLGVLVSFPAKHGNA 573
Qy 564 HRLLELLELPELFLGLGVLVFLVLYKWLVCWAARAAAPSILIHFINFLSHSPNRL 623
Db 574 LDRIQVIFPMIFLNSFLGYLALLILIKW-----CRGSOADLYHVMYMF-----DPSGNL 625
Qy 624 ----LYPROVVQATLVVLALAMVPILLGLTPLLHLILHRRRLRRRPPADQENKAGLLD 679
Db 626 GENQLFWQKEQLILLMAIIVAPWMLFPKPFILKLNERPQ-----CHTYRFLGTSE 680
Qy 680 L-PDASVNGSSDEKAGLDEEAEALVPSSEVLVHQHAIHTIEFCLGCVSNTASVLRWA 738
Db 681 MDPD-----SEPDARSRHDDFNF-----SEVFHQMIHSIEFVLGAVSNTASVLRWA 729
Qy 739 LSLAHQALSEVLWAMVMRIGLGLGREGVGAUVLPIFAAFVMTVAIILVMEGLSAPLH 798
Db 730 LSLAHSELSVTFYKLLVLAWGDNLV-VKLVGLV-----IFGFATAFILLGMSLSAPLH 784
Qy 799 ALRLHWVEFQNFYSGTYKLSPP 822
Db 785 ALRLHWVEFMNFYHGDYKFRPF 808

RESULT 5

US-10-449-902-41603
; Sequence 41603, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 41603
; LENGTH: 923
; TYPE: PRN
; ORGANISM: Oryza sativa
US-10-449-902-41603

Query Match 26.1%; Score 1113; DB 6; Length 923;
Best Local Similarity 31.3%; Pred. No. 1.5e-82;
Matches 304; Conservative 143; Mismatches 307; Indels 218; Gaps 33;
Qy 3 SMFRESEVALVOLFLPTAAAYTCVSRGLGELVFEFDLNA-----SVSAFORRVV 53
Db 10 SLFRSRPMYSVITAMTDIAYDTIRELGNFLGFHVVDLNAAGKASVNDKNLAGYKKR--- 66
Qy 54 DWRCCELEKTFYFQESVRRAGLVLPKGRLPAPPPDLRIIOETERLAQE---IRD 110
Db 67 -VQDCIMYERKLUHDEQIK-----EYNIPAPPP---LREHEERHHTKONAVLPD 113
Qy 111 VRGNOQALRAQLHQ-LQLHAALVR-----OQHEPQ-----L 140
Db 114 IRAHLEPLEERLRSVQVQADITRAISELTERVWYKTCRNQVGRVEAHDDEKHEVPYSA 173
Qy 141 AAHTDGA-----SER---TFLQAPGPGHODLRV 167
Db 174 AAAAPEGALVPGTTAAGLPAALAPPKQKNDVGLGERPVGTPPGTTGGG---DSKL 230
Qy 168 N-----FV-----AGAVEPHKAPALERLLWRACRGFLIA 196
Db 231 NQLGVSLISADHAPSAGGAAGKVLKPFRTLLTGVLPTDKSSFFQRLYRISGNALT 290
Qy 197 SFRELEQPLEHPVTGEPATMTTF-LISYNGEQIGQKIRKITDCFHCHVFPFLQBEARLG 255

Db 291 HTFDIPVDPVAGSELVNKSVFSLIIVVSELEKRIKVKQCFQATIVRMDPTQAAIDE 350
Qy 256 ALQOI-QOQSQELQ-----EVLGETERFLSQVLRVLIQLLP--PGQVQVHKMKAVYALNQC 309
Db 351 ALHRTEAQIDEQMLKVNTRLRDIEQALRQLAGDPQNVSPLEWQVLLRCEKAVCDTMMKC 410
Qy 310 SVSTTHKCLIA-EAMCSVRDLPALQEAR-----DSSMEBGVSVAHRIPCRD--MPPTL 361
Db 411 HFLYLT---MVAPEGCPVEEVDNLKIAVKRAVPPGDKIPEPAVEVAPEKPIRDPGPPPY 467
Qy 362 INTNFTASFOGIVDRYGVGRYQEVNPPAPYTIITFPFLFVNFVGVGVGHLLMFLPALAMV 421
Db 468 FKLNFTSIFQIVNTYGVPRYQEVNPGFTIIVTFPFLFGVMYGDIGHGTFVTIISLLMI 527
Qy 422 LAENR-PAVKAA--ONEIWQTFRRYLLLLMGLFSIYTGFIYNECFSRATSIFFSGMSV 478
Db 528 IPEDRLTRLKESGNMNEIFRMCFCGRYLLVCMVMGIYCGTLYNDCMSIPVSLYAPTWEY 587
Qy 479 AAMANQSGSDAFLAQTMLTLDPNVTGVFLGPPYFGIDPIWLSLAANHLSPLNSPKMKMS 538
Db 588 EGTADIPA-----TRIGAV-----YPYGVDPAWYHTKNQTLTFNGMKMKTS 628
Qy 539 VILGVVHMAFGVVLGVFNHVFQGRHLLLELTELTLGLGLFVFLVIVKMLCVAA 598
Db 629 VILGVQWTFGLIINGLLNHHPRDOISIWLEFLPQLLPMCTGFWCMLIIIIKWCDSQ 688
Qy 599 RAASPSILIH-FINMPL-FSHSPSNRLLYPRQEVVQATLVVLALAMVPILLGLTPLLHLL- 655
Db 689 TSQEPNLIQTMITMFLGLGVKEGAQLYAGQSTVQAILLIIFAGSIPVMLLAKPCILAY 748
Qy 656 -----HRHRRRLRRRADRQENKAGLLDLPDASVNGSSDEKAGGLDDEEAEALVP 708
Db 749 CSKDDHGHGAKPQAAQPA-----AAAAAG-----DVEAKGFD--EKGEVIP 787
Qy 709 SEV-----LMHQAIHTIEFCLGCVSNTASYLRMLWLSLAHQAOLSEVL 750
Db 788 AEAPRAGAGAAHEEHEHTFGDYMIHQSIHTIEFVLGTVTNTASYLRMLWLSLAHSELDYF 847
Qy 751 WAMVMRIGLGLGREGVGAUVLPIFAAFVMTVAIILVMEGLSAPLHRLHWVEFQNK 810
Db 848 WTKLI-LEYGMYKNSAFFSMVAV---AVMFMVTAGVLLSMDVLECFHLALRLHWVEFQNK 903
Qy 811 FVSGTYGKLSPP 822
Db 904 FFHADGYAFKPF 915

RESULT 6

US-10-449-902-52773
; Sequence 52773, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 52773
; LENGTH: 489
; TYPE: PRN
; ORGANISM: Oryza sativa
US-10-449-902-52773

[illegible]

```

RESULT 7
US-10-449-902-29928
; Sequence 29928, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205V1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29928
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-29928

Query Match      14.7%; Score 627.5; DB 6; Length 357;
Best Local Similarity 42.1%; Pred. No. 1.4e-43;
Matches 164; Conservative 44; Mismatches 129; Indels 53; Gaps 15;

QY      441  FRGRYLLLLMGLFSIYGFYNECFSRATSIPPSGWSVAAWANGSGWSDAFLAQHTMLT 500
Db      3    FGGRYVLLMALFSYGLLIYNEFFSPHFIFGKS---AYCREKTCSDA----HT----- 51

```

| | | | |
|----|-----|---|-----|
| Qy | 501 | DPNVTGVF--LGPYPFGIDPTWSLAANHLSFLNSFKMKMSVILGVVHMAFGVVVLGVFNHV | 558 |
| | | : : | : |
| Db | 52 | ----AGLIKVPDPYFPFGVDPSPWRGSRSELPFLNLSLKMKMSIIMGVTTQMNLGITVLSYFDAK | 107 |
| | | : : | : |
| Qy | 559 | HFGQRHRLLETLPELTLLGLFGVLVLYVKWLCVWAARAASPSILIH-FINMPLFSH | 617 |
| | | : : | : |
| Db | 108 | FHGNALDIRYQIPQWIFLNSFGYLALILIKW-C-----TGSQADLVHWIMYMF---- | 158 |
| | | : : | : |
| Qy | 618 | SPSNRL-----LYPRQBVVQATLVVLALAMVPILLLGTPHLHLHRHRRRLRRPADRQEEN | 673 |
| | | : : | : |
| Db | 159 | DPSGNLGENQLFWGQKELQIILLMLAIVAPWMPLPKPFILKKLKERFQ-----GHTYR | 213 |
| | | : : | : |
| Qy | 674 | KAGILDL- PDASVNGWSDEKAGGLDDEBEAEALVPSEVLMMQAHTTFCLGCVSNTAS | 732 |
| | | : : | : |
| Db | 214 | FLGTSEMDDP-----SEPDSARSRHDDNF-----SEVFHQMTHSIEFVLGAVSNTAS | 262 |
| | | : : | : |
| Qy | 733 | YLRWLALSIAHAQLSEVLWAMVMRMICLGLGREVGVAADVLPFAAFAMTVAILLMVBEG | 792 |
| | | : : | : |
| Db | 263 | YLRWLALSIAHSELSTVFEYKLLVLAWGYDNLV-VKLVLV-----IFSPTAFILLGMES | 317 |
| | | : : | : |
| Qy | 793 | LSAFHALRLHWVEFQNKFYSGTGKYKLSPP | 822 |
| | | : : | : |
| Db | 318 | LSAFHALRLHWBFBMNKFYHGDGYKFRPF | 347 |
| | | : : | : |

RESULT 8
US-10-953-349-18079
; Sequence 18079, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POL
; TITLE OF INVENTION: ENCODED THERBY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18079
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-18079

| | | | | | |
|-----------------------|--------|--|-------|-------------|-----|
| Query Match | 14.1% | Score 600; | DB 6; | Length 314; | |
| Best Local Similarity | 44.7%; | Pred. No.2e-41; | | | |
| Matches | 142; | Conservative | 39; | Mismatches | 99; |
| | | | | Indels | 38; |
| | | | | Gaps | 8; |
| | | | | | |
| Qy | 512 | YPFGIDPISLAANHLSFLNSFKMKMSVILGVHMAFGVLGVFNHVFHQGRHLLLETL | 571 | | |
| Db | 19 | YPFGVDPVHGTSGELPFLNSLKKMKMSILLGVAQNLGIVMSYFNAIFRNSVNVWFQFI | 78 | | |
| | | | | | |
| Qy | 572 | PELTFLGLGFLVFLVIYKWLCVMAARAAPSILIH-FINMFLFSHSPSNRL---- | 626 | | |
| Db | 79 | PQMIFLNSLFGYLSLLIIVKM-----ATGSOADYLHILYMF-----SPTDDLGENQLFA | 129 | | |
| | | | | | |
| Qy | 627 | ROEVVQATLVVLAMVPILLGTPLHLHRRRLRRRADROEENKAGLLDLPDASVN | 686 | | |
| Db | 130 | GQKNLQVLVLLLAIVISVPWMLLPFFILKKQHEAR-----HGVEASYPL----- | 173 | | |
| | | | | | |
| Qy | 687 | GWSSDE--EKAGGLDDEEAEELVPSEVLHMQAIIHTIEFLGCVSNTASYLRWLWLSLAHA | 744 | | |
| Db | 174 | -QSTDESLOESNHDSHGHEEFSEFVFGVLIHTIEFVLGAVSNTASYLRWLWLSLAHS | 232 | | |
| | | | | | |
| Qy | 745 | QLSEVLWAMVMRIGLGLGREVGVAAVLVPTFAAFVAMTVAILLWMEGLSAFLHALRLHW | 804 | | |
| Db | 233 | ELSSVFEKVLMMAGVNNVILLVGLVIFFA-----TVGVLLWMTLSAFLHALRLHW | 287 | | |
| | | | | | |
| Qy | 805 | VEFONKFYSGTGYKLSPP | 822 | | |
| Db | 288 | VEFONKFYEGDGYKFFHP | 305 | | |

[illegible]

```
RESULT 9
US-10-953-349-18080
; Sequence 18080, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18080
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-18080

Query Match      12.2%; Score 521; DB 6; Length 273;
Best Local Similarity 43.7%; Pred. No. 4.6e-35;
Matches 129; Conservative 36; Mismatches 92; Indels 38; Gaps 8;

QY 535 MMSVILGVVHMAFGVVLGVFNHVFQGRHRLLELTPELTLLGLFGYLVFLVIYKWL 594
Db 1 MMSVILGVVHMAFGVVLGVFNHVFQGRHRLLELTPELTLLGLFGYLVFLVIYKWL 58
QY 595 VMAAASPSILIH-FINMFLFSHSPNRL-----LYPRQEVVQATLVVLALAMVPILLG 649
Db 59 ----ATGSGADLYHILYMFEL---SPTDDLGENQLFAGQKNLQLVLLLVISVPMLLP 111
QY 650 TPLHLHRRHRRRLRRPADRQENKAGLLDLDPASVNGWSSDE--EKAGGLDDEEAEIV 707
Db 112 KPFLKKQHEAR-----HGVSAYPL-----QSTDESLOVESNHDHSHGHEBFE 154
QY 708 PSEVLHQAHTTIEFCIGCVSNATSYLRMLWALSLAHQAQLSEVLWAMVNRIGLGLGREVG 767
Db 155 PSEVPVQLHTTIEFVLGAVSNATSYLRMLWALSLAHSELSSVFYEKVLWMAWYNNVIL 214
QY 768 AAVLVPIFAAPAVMTVAILLVMEGLSAFLHRLHWHVEFQNKFGYSGTGKLSPF 822
Db 215 IVGLVFIFA-----TVGVLLVMTLSAFLHRLHWHVEFQNKFGYSGTGKLSPF 264

RESULT 10
US-10-953-349-18081
; Sequence 18081, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 18081
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-18081

Query Match      12.0%; Score 511; DB 6; Length 271;
Best Local Similarity 43.3%; Pred. No. 3e-34;
Matches 127; Conservative 36; Mismatches 92; Indels 38; Gaps 8;

QY 537 MSVILGVVHMAFGVVLGVFNHVFQGRHRLLELTPELTLLGLFGYLVFLVIYKWL 596
Db 1 MSVILGVVHMAFGVVLGVFNHVFQGRHRLLELTPELTLLGLFGYLVFLVIYKWL 56
QY 597 AARAASPSILIH-FINMFLFSHSPNRL-----LYPRQEVVQATLVVLALAMVPILLGTP 651
```

```
Db 57 --ATGSGADLYHILYMFEL---SPTDDLGENQLFAGQKNLQLVLLLVISVPMLLP 111
QY 652 LHLHRRHRRRLRRPADRQENKAGLLDLDPASVNGWSSDE--EKAGGLDDEEAEIVPS 709
Db 112 FILKKQHEAR-----HGVSAYPL-----QSTDESLOVESNHDHSHGHEBFEPS 154
QY 710 EVLHQAHTTIEFCIGCVSNATSYLRMLWALSLAHQAQLSEVLWAMVNRIGLGLGREVG 769
Db 155 EVFVHQLHTTIEFVLGAVSNATSYLRMLWALSLAHSELSSVFYEKVLWMAWYNNVIL 214
QY 770 VVLVPIFAAPAVMTVAILLVMEGLSAFLHRLHWHVEFQNKFGYSGTGKLSPF 822
Db 215 GLIVFIFA-----TVGVLLVMTLSAFLHRLHWHVEFQNKFGYSGTGKLSPF 262

RESULT 11
US-10-953-349-30245
; Sequence 30245, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 30245
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-953-349-30245

Query Match      8.5%; Score 361; DB 6; Length 306;
Best Local Similarity 29.9%; Pred. No. 5.8e-22;
Matches 87; Conservative 63; Mismatches 101; Indels 40; Gaps 5;

QY 4 MFRSEVALVQLFPTAAAYTCVSRGLBELGLVFRFDLNASVAFORFVVDVWRCEELEK 63
Db 3 LMRSEAMQLQVIPTESAHIAVSHLGLDGLIQFDNLNADKSPFORTYAAQIKRCAEMAR 62
QY 64 TPTFLOEBVRAGLVLPKGRLPAPPDRDLRIQEETERLAQELRDYRGNQOALRAQLH 123
Db 63 KLRFFKEQMSKAGILVSPMQS---TETPLDFFDMVKUGLEAEUTEVYNADEKLQRAHN 119
QY 124 QLQLHAAVLRQHE-----PQLAAHTDGAERTPLIQAPEGPHQDL----- 165
Db 120 ELLEYSTVLQKAGEFFYSQAQSAQAHRQMEANQSGETSLESPLLE-----QDMLTDAS 173
QY 166 ---RVNVFVAGAVEPHKAPALBRLLWRACRGPLIASFRELEQPLEHPVTGEPATWMTPLIS 222
Db 174 KQVKGLSGLVLPKEKAMAFERILFRATRGNILLRQESVDEPVDTPQSGEKYKNTFVVF 233
QY 223 YWGEIQGQIRKRTDCCHCHVFPEFLQGEERLALGALQQLQQQSQELQEVLGE 273
Db 234 YSGERAKAKILKICDAFRANKYFP-----PEDLAKQTHVTQEVSGK 274

RESULT 12
US-10-953-349-30244
; Sequence 30244, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
```



```

Db 495 RVC-----RLRANR-----VLRACERAVKSNACYWAVLLVFLNTLTIASHHQ 542
QY 412 LMFPLALAMVLAENPAVKAONEIWOTFFRGYLLLLMGLF-SIYTGFIYN--ECFSRA 468
Db 543 PVWL-----TQIOEYANKVLLCLFTVEMLLKLYGLGPSAYVSSFFNRFDCFVVC 591
QY 469 TSIF-----PSGWSVAAMANQSGMSDAFLAHTMLTLDPNVTGVLGYPFGID 517
Db 592 GGILETTLVEVGAMPGLGISVLRV-----LLRI-----FKVT 625
QY 518 PIWSLAANHL-SFLNSFKMKMSVIL-----GVVHMAFGVVL--GVENHVHFGQRH--RL 567
Db 626 RHWASLSNLVASLNSMKSIASILLLLPLFIIFSLGMLFGGKEN---PDQTHKRST 682
QY 568 LETLPE--LTFLGLFGYLVFLVVKWLCVWAARAAPSIL--IHPINMFLFSHSPSNRL 623
Db 683 FDTFFQALLTVQILTGEDWNVYDGINMAYGG-PFFPGMLVCIYFIILFIC---GNYI 737
QY 624 LYPRQEVVQATLVVLALAM-----VPILLGTPLHLLHRRH 660
Db 738 LLN-----VFLAIVDNLASDAGTAKDKGKSKNEKDLPOENEGLVFGVEKEEEE 788
QY 661 RLRRPADRQENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLHMQAIHTI 720
Db 789 GARREGADMEEBEE-----EEEEEEEEEGAGGV--ELLQEVVPKVKVPIPEGSA 839
QY 721 BECLGCVNTASYLRLWALSIAHQL--SEVLWAMVMRIGLGLGREVGVA-----VVLV 773
Db 840 FFCL-----SQTNPLRKGCHTLHHVFTNLILVFIILSSVSLAEDPIRAHSFRNHILGY 895
QY 774 PIFAAPAVMTVAILLVMEGLSAPLH 798
Db 896 FDYAFTSIFTVEILLKMTVEGAPLH 920

```

Search completed: June 29, 2006, 13:13:30
 Job time : 23 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2006, 13:13:20 ; Search time 194 Seconds
(without alignments)
1937.278 Million cell updates/sec

Title: US-10-783-519-2.

Perfect score: 822

Sequence: 1 MGSMPRSEVALVOLFLPRA.....HWVEFQNKFGYGYKLSPF 822

Scoring table:

Capop 60.0 , Gapext 60.0

Searched: 2589679 seqs, 457216429 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2589342

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq.8.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*
- 9: Geneseq2005s.*
- 10: Geneseq2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 822 | 100.0 | 822 | 2 AAW41943 | AAW41943 Human OC- |
| 2 | 822 | 100.0 | 822 | 5 ABB81811 | Abb81811 Human OC- |
| 3 | 822 | 100.0 | 822 | 8 ADR44119 | Adr44119 Human 116 |
| 4 | 822 | 100.0 | 822 | 9 ADY92606 | Ady92606 Human 116 |
| 5 | 545 | 66.3 | 830 | 8 ADI28834 | Adi28834 Human mod |
| 6 | 545 | 66.3 | 830 | 8 ADI28834 | Adi28834 Human mod |
| 7 | 545 | 66.3 | 830 | 8 ADI28834 | Adi28834 Human mod |
| 8 | 501 | 60.9 | 787 | 8 ABM81281 | Abm81281 Tumour-as |
| 9 | 445 | 54.1 | 787 | 8 ABM84820 | Abm84820 Human dia |
| 10 | 385 | 46.8 | 614 | 2 AAW98112 | Aaw98112 T-cell me |
| 11 | 385 | 46.8 | 614 | 8 ABM81282 | Abm81282 Tumour-as |
| 12 | 325 | 39.5 | 614 | 2 AAW98113 | Aaw98113 T-cell me |
| 13 | 325 | 39.5 | 614 | 6 ABR43098 | AbR43098 Human T-c |
| 14 | 291 | 35.4 | 643 | 5 ABP41524 | Abp41524 Human ova |
| 15 | 124 | 15.1 | 124 | 7 ADD27652 | Add27652 Human adi |
| 16 | 112 | 13.6 | 203 | 7 ADD27225 | Add27225 Human adi |
| 17 | 34 | 4.1 | 60 | 5 ADK35292 | Adk35292 Novel hum |
| 18 | 33 | 4.0 | 33 | 8 ADR44126 | Adr44126 Human OC- |
| 19 | 33 | 4.0 | 33 | 9 ADY92613 | Ady92613 Human 116 |
| 20 | 32 | 3.9 | 32 | 8 ADR44121 | Adr44121 Human OC- |
| 21 | 32 | 3.9 | 32 | 8 ADY92608 | Ady92608 Human 116 |
| 22 | 30 | 3.6 | 30 | 8 ADR44124 | Adr44124 Human OC- |
| 23 | 30 | 3.6 | 30 | 9 ADY92611 | Ady92611 Human 116 |

| | | | | | | |
|----|----|-----|-----|---|----------|--------------------|
| 24 | 29 | 3.5 | 29 | 6 | ABR43100 | AbR43100 Human T-c |
| 25 | 27 | 3.3 | 830 | 8 | ADH09947 | Adh09947 Human hos |
| 26 | 27 | 3.3 | 831 | 5 | ABB08457 | Abb08457 Human tum |
| 27 | 27 | 3.3 | 831 | 8 | ADH09948 | Adh09948 Human hos |
| 28 | 27 | 3.3 | 831 | 8 | ADH09952 | Adh09952 Human hos |
| 29 | 27 | 3.3 | 831 | 8 | ADH09953 | Adh09953 Human hos |
| 30 | 27 | 3.3 | 831 | 8 | ADI28837 | Adi28837 Human mod |
| 31 | 26 | 3.2 | 26 | 8 | ADR44122 | Adr44122 Human OC- |
| 32 | 26 | 3.2 | 26 | 9 | ADY92609 | Ady92609 Human 116 |
| 33 | 24 | 2.9 | 24 | 6 | ABR43102 | AbR43102 Human T-c |
| 34 | 24 | 2.9 | 24 | 6 | AAE37759 | Aae37759 TIRC7 pep |
| 35 | 24 | 2.9 | 24 | 6 | AAE37809 | Aae37809 TIRC7 pep |
| 36 | 24 | 2.9 | 513 | 6 | ADA54643 | Ada54643 Human pro |
| 37 | 24 | 2.9 | 558 | 4 | ABG09380 | Abg09380 Novel hum |
| 38 | 24 | 2.9 | 821 | 7 | ADC87337 | Adc87337 Human GPC |
| 39 | 24 | 2.9 | 840 | 5 | ABP69796 | Abp69796 Human pol |
| 40 | 24 | 2.9 | 840 | 8 | ADH09949 | Adh09949 Human hos |
| 41 | 24 | 2.9 | 840 | 8 | ADI28836 | Adi28836 Human mod |
| 42 | 24 | 2.9 | 841 | 4 | ABG60100 | Abg60100 Human tra |
| 43 | 24 | 2.9 | 847 | 8 | ADR44120 | Adr44120 Human OC- |
| 44 | 24 | 2.9 | 847 | 9 | ADY92607 | Ady92607 116 kd os |
| 45 | 23 | 2.8 | 23 | 8 | ADR44125 | Adr44125 Human OC- |

ALIGNMENTS

RESULT 1

AAW41943

ID AAW41943 standard; protein; 822 AA.

XX AAW41943;

XX 02-JUL-1998 (first entry)

XX Human OC-116 kDa protein.

XX Osteoclast; human; OC-116 kDa; screening; cell surface marker; probe.

XX Homo sapiens.

XX WO9803651-AL.

XX 29-JAN-1998.

XX 10-JUL-1997; 97WO-US012569.

XX 19-JUL-1996; 96US-00684932.

XX (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.

XX Stashenko P, Li Y, Wucherpfennig AL;

XX WPI; 1998-120776/11.

XX N-PSDB; AAV04529.

XX Osteoclast specific or related DNA sequence - useful as probe to screen genomic DNA or cDNA library, or as osteoclast cell surface marker.

XX Example 8; Page 61-64; 75pp; English.

XX The present sequence represents the human OC-116 kDa protein sequence which is used in an example of the present invention, which describes isolated osteoclast specific or related DNA sequences. The present invention also describes: a DNA construct capable of replicating and optionally expressing, in a host cell, osteoclast specific or related DNA, comprising an osteoclast specific or related DNA sequence and CC sequences necessary for transforming or transfecting a host cell, and for replicating and optionally expressing an osteoclast specific or related CC DNA sequence in a host cell; and a cell stably transformed or transfected CC with the DNA construct. The osteoclast specific or related DNA sequence CC can be used as a probe to screen a genomic DNA or cDNA library for osteoclast specific or related DNA sequences, or as a osteoclast cell

```
CC surface marker
XX
SQ Sequence 822 AA;

Query Match 100.0%; Score: 822; DB 2; Length 822;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSMFRSEVALVQLFLPTAAAYTCVSRIGELGLVEFRDLNASVSFAQRRFVVDVWRCSE 60
DB 1 MGSMFRSEVALVQLFLPTAAAYTCVSRIGELGLVEFRDLNASVSFAQRRFVVDVWRCSE 60

QY 61 LEKTTFTLQEEVRRAGLVLPKPKGRLPAPPDRLRIQETELRLAQELRDVRGNQOALRA 120
DB 61 LEKTTFTLQEEVRRAGLVLPKPKGRLPAPPDRLRIQETELRLAQELRDVRGNQOALRA 120

QY 121 QLHQQLHAAVLRQGHPEPQLAAHTDGASERTPLQAQGGPHQDLRVNPFVAGAVEPHKAP 180
DB 121 QLHQQLHAAVLRQGHPEPQLAAHTDGASERTPLQAQGGPHQDLRVNPFVAGAVEPHKAP 180

QY 181 ALERLLWRACRGFLTASPRELEQPLEHPVTGEPATWMTFLISYWGEOIQGKIRKIDTDPH 240
DB 181 ALERLLWRACRGFLTASPRELEQPLEHPVTGEPATWMTFLISYWGEOIQGKIRKIDTDPH 240

QY 241 CHVFPFLQEEVRRAGLVLPKPKGRLPAPPDRLRIQETELRLAQELRDVRGNQOALRA 300
DB 241 CHVFPFLQEEVRRAGLVLPKPKGRLPAPPDRLRIQETELRLAQELRDVRGNQOALRA 300

QY 301 AVYLAHQCSVSTTHKCLIAEAWCSVRDLPAQELRDSSMEEGVSVAHRIPCRDMPT 360
DB 301 AVYLAHQCSVSTTHKCLIAEAWCSVRDLPAQELRDSSMEEGVSVAHRIPCRDMPT 360

QY 361 LIIRNRPFTASFOGIVDRYGVGRYQEVNPAFYIITFPFLFAVMFGDVGHGLMFLPALAM 420
DB 361 LIIRNRPFTASFOGIVDRYGVGRYQEVNPAFYIITFPFLFAVMFGDVGHGLMFLPALAM 420

QY 421 VLAENRPVAKAAQNEIWTQFFRGRYLLMLLGLFSIYTGFIYNECFSRATSIIPSGMSVAA 480
DB 421 VLAENRPVAKAAQNEIWTQFFRGRYLLMLLGLFSIYTGFIYNECFSRATSIIPSGMSVAA 480

QY 481 MANQSGWSDAFLAQHTMLTLDNVTGVFLGPPYFGIDPIWSLAANHLSFLNSFKMKMSVI 540
DB 481 MANQSGWSDAFLAQHTMLTLDNVTGVFLGPPYFGIDPIWSLAANHLSFLNSFKMKMSVI 540

QY 541 LGVVMHAFGVVLGVFNHVFHQRHLLLETLPELTLGLFGVLPVLYKWLVCVWAARA 600
DB 541 LGVVMHAFGVVLGVFNHVFHQRHLLLETLPELTLGLFGVLPVLYKWLVCVWAARA 600

QY 601 ASPSILIHFINFLFSHSPSNRLLYPRQEVQATLVVLALAMVPIILLGTPLHLHRHR 660
DB 601 ASPSILIHFINFLFSHSPSNRLLYPRQEVQATLVVLALAMVPIILLGTPLHLHRHR 660

QY 661 RLRRPADQENKAGLLDLPASVNGWSSDEKAGGLDDEAEELVPSEVLHQAIHTI 720
DB 661 RLRRPADQENKAGLLDLPASVNGWSSDEKAGGLDDEAEELVPSEVLHQAIHTI 720

QY 721 EFCLGCVNSTASYLRWLWLSLAHAQLSEVLWAMVMRIGLIGREVCVAAVLPVFAFA 780
DB 721 EFCLGCVNSTASYLRWLWLSLAHAQLSEVLWAMVMRIGLIGREVCVAAVLPVFAFA 780

QY 781 VMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYSGTGKLSPPF 822
DB 781 VMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYSGTGKLSPPF 822

RESULT 2
ABB81811
ID ABB81811 standard; protein; 822 AA.
XX
AC ABB81811;
XX
XX 16-SEP-2002 (first entry)
XX
```

```
DE Human OC-116KDa.
XX
XX Human; osteoclast; gene therapy; aberrant bone resorption; OC-116KDa.
XX Homo sapiens.
XX US6403304-B1.
XX
XX 11-JUN-2002.
XX 19-JUL-1996; 96US-00684932.
XX
XX 06-APR-1993; 93US-00045270.
XX 23-FEB-1995; 95US-00392678.
XX 20-JUL-1995; 95US-0001292P.
XX 22-FEB-1996; 96US-00605378.
XX
XX (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.
XX
XX Stashenko P, Li Y, Wuchertpfennig AL;
XX WPI; 2002-536031/57.
XX N-PSDB; ABN86735.
XX
XX Novel isolated osteoclast-specific or -related DNA sequence, useful for
XX producing gene products useful in the therapeutic treatment or diagnosis
XX of disorders involving aberrant bone resorption.
XX
XX Example 8; Fig 3; 34pp; English.
XX
XX The invention relates to novel human osteoclast-specific or -related cDNA
XX sequences. The sequence represents human osteoclast 116KDa (OC-116KDa).
XX The sequences may have a use in gene therapy. The sequences of the
XX invention are useful in the production of gene products useful in the
XX therapeutic treatment or diagnosis of disorders involving aberrant bone
XX resorption, for generating peptides which are useful for producing
XX antibodies for identifying osteoclast-specific or -related peptides or
XX gene products
XX
XX Sequence 822 AA;

Query Match 100.0%; Score 822; DB 5; Length 822;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSMFRSEVALVQLFLPTAAAYTCVSRIGELGLVEFRDLNASVSFAQRRFVVDVWRCSE 60
DB 1 MGSMFRSEVALVQLFLPTAAAYTCVSRIGELGLVEFRDLNASVSFAQRRFVVDVWRCSE 60

QY 61 LEKTTFTLQEEVRRAGLVLPKPKGRLPAPPDRLRIQETELRLAQELRDVRGNQOALRA 120
DB 61 LEKTTFTLQEEVRRAGLVLPKPKGRLPAPPDRLRIQETELRLAQELRDVRGNQOALRA 120

QY 121 QLHQQLHAAVLRQGHPEPQLAAHTDGASERTPLQAQGGPHQDLRVNPFVAGAVEPHKAP 180
DB 121 QLHQQLHAAVLRQGHPEPQLAAHTDGASERTPLQAQGGPHQDLRVNPFVAGAVEPHKAP 180

QY 181 ALERLLWRACRGFLTASPRELEQPLEHPVTGEPATWMTFLISYWGEOIQGKIRKIDTDPH 240
DB 181 ALERLLWRACRGFLTASPRELEQPLEHPVTGEPATWMTFLISYWGEOIQGKIRKIDTDPH 240

QY 241 CHVFPFLQEEVRRAGLVLPKPKGRLPAPPDRLRIQETELRLAQELRDVRGNQOALRA 300
DB 241 CHVFPFLQEEVRRAGLVLPKPKGRLPAPPDRLRIQETELRLAQELRDVRGNQOALRA 300

QY 301 AVYLAHQCSVSTTHKCLIAEAWCSVRDLPAQELRDSSMEEGVSVAHRIPCRDMPT 360
DB 301 AVYLAHQCSVSTTHKCLIAEAWCSVRDLPAQELRDSSMEEGVSVAHRIPCRDMPT 360

QY 361 LIIRNRPFTASFOGIVDRYGVGRYQEVNPAFYIITFPFLFAVMFGDVGHGLMFLPALAM 420
DB 361 LIIRNRPFTASFOGIVDRYGVGRYQEVNPAFYIITFPFLFAVMFGDVGHGLMFLPALAM 420
```

QY 421 VLAENRPAVKAQNEIWTQTFRGRYLLMLGLPSIYTGFIYNECFSRATSIFFSGWSVAA 480
DB 421 VLAENRPAVKAQNEIWTQTFRGRYLLMLGLPSIYTGFIYNECFSRATSIFFSGWSVAA 480
QY 481 MANQSGWSDAFLAQHTMTLDPNVTGVFLGPYPFGIDPIWISLAANHLSFLNSFKMKMSVI 540
DB 481 MANQSGWSDAFLAQHTMTLDPNVTGVFLGPYPFGIDPIWISLAANHLSFLNSFKMKMSVI 540
QY 541 LGVVMHAFGVVLGVFNHVFQGRHRLLETLPELTFLGLFGYLVFLVIYKWLVCWAAARA 600
DB 541 LGVVMHAFGVVLGVFNHVFQGRHRLLETLPELTFLGLFGYLVFLVIYKWLVCWAAARA 600
QY 601 ASPSILIHFINMFLFSHPNRLLYPROEVVQATLVVLALAMPVILLGLTPLHLHRRR 660
DB 601 ASPSILIHFINMFLFSHPNRLLYPROEVVQATLVVLALAMPVILLGLTPLHLHRRR 660
QY 661 RLRRRPAQREKAGLLDLPDASVNGWSDEEKAGGLDDEEAELVPSEVLHQAIHTI 720
DB 661 RLRRRPAQREKAGLLDLPDASVNGWSDEEKAGGLDDEEAELVPSEVLHQAIHTI 720
QY 721 EFCGCVSNTASYRLRWALSLAHQAQSEVLWAMVMRIGLGLGREGVGAADVLPPIFAAFA 780
DB 721 EFCGCVSNTASYRLRWALSLAHQAQSEVLWAMVMRIGLGLGREGVGAADVLPPIFAAFA 780
QY 781 VMTVAILLVMEGLSAFLHRLHWHVEFQNKFSYGTGYKLSPP 822
DB 781 VMTVAILLVMEGLSAFLHRLHWHVEFQNKFSYGTGYKLSPP 822

RESULT 3

ADR44119
ID ADR44119 standard; protein; 822 AA.

XX AC ADR44119;

XX DT 04-NOV-2004 (first entry)

XX DE Human 116-kDa osteoclast proton pump subunit (OC-116 kDa) protein.

XX KW Human; OC-116 kD; 116-kDa osteoclast proton pump subunit; therapy;
bone degradation; osteoporosis; osteoarthritis.

XX OS Homo sapiens.

| Key | Location/Qualifiers |
|-----------|------------------------------|
| FT Domain | 319..422 |
| FT Domain | /note = Transmembrane domain |
| FT Domain | 438..463 |
| FT Domain | /note = Transmembrane domain |
| FT Domain | 537..558 |
| FT Domain | /note = Transmembrane domain |
| FT Domain | 571..600 |
| FT Domain | /note = Transmembrane domain |
| FT Domain | 632..653 |
| FT Domain | /note = Transmembrane domain |
| FT Domain | 764..796 |
| FT Domain | /note = Transmembrane domain |

XX PN US6777537-B1.

XX PD 17-AUG-2004.

XX PF 18-JUL-2000; 2000US-00618304.

XX PR 22-FEB-1996; 96US-00605378.

XX PA (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.

XX PI Stashenko P, Li Y;

XX DR WPI; 2004-591304/57.

XX DR N-PSDB; ADR44118.

XX

PT New isolated human 116-kDa osteoclast (OC-116 kDa) proton pump subunit
polypeptide, useful for preventing and/or treating diseases with bone
degradation, such as osteoporosis and osteoarthritis.

XX Claim 2; SEQ ID NO 2; 16pp; English.

XX The present invention relates to a human 116-kDa osteoclast proton pump
subunit (OC-116 kDa) polypeptide and its encoding polynucleotide. The
invention is useful for the prevention and treatment of diseases or
conditions associated with aberrant expression or activity of the OC-116
kDa protein and bone degradation, such as osteoporosis and
osteoarthritis. The present sequence is human OC-116 kDa protein.

XX Sequence 822 AA;

Query Match 100.0%; Score 822; DB 8; Length 822;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSMPRSEEEVALVQLFLPTAAAYTCVSRIGELGLVEFRDLNANVSAPQRRFVVDVWRCCE 60

DB 1 MGSMPRSEEEVALVQLFLPTAAAYTCVSRIGELGLVEFRDLNANVSAPQRRFVVDVWRCCE 60

QY 61 LEKTFPLEQEVRRAGLVLPKGRLPAPPDRLLRIQETERLAQELDRVGRNQOALRA 120

DB 61 LEKTFPLEQEVRRAGLVLPKGRLPAPPDRLLRIQETERLAQELDRVGRNQOALRA 120

QY 121 QLHQLQHAALVLRQGHQEPQLAAHTDGASERTPLQAPGGPHQDLRVNFVAGAVEPHKAP 180

DB 121 QLHQLQHAALVLRQGHQEPQLAAHTDGASERTPLQAPGGPHQDLRVNFVAGAVEPHKAP 180

QY 181 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGSGIQKIRKITDCPH 240

DB 181 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGSGIQKIRKITDCPH 240

QY 241 CHVFPFLOQBEARLQALQQLQQSQSLEQVLGTERFLSQVLGRVLQQLPPGVQVHKMK 300

DB 241 CHVFPFLOQBEARLQALQQLQQSQSLEQVLGTERFLSQVLGRVLQQLPPGVQVHKMK 300

QY 301 AVYALNQCNSVSTTHKCLIAEAWCSVRDLPALQEARLDSMEGVSVAHRIFCRDMPT 360

DB 301 AVYALNQCNSVSTTHKCLIAEAWCSVRDLPALQEARLDSMEGVSVAHRIFCRDMPT 360

QY 361 LIRTNRTASFOGIVDRYGVRYQEVNPAPYTIITPPFLFVAVMGDVGHLMLFLALAM 420

DB 361 LIRTNRTASFOGIVDRYGVRYQEVNPAPYTIITPPFLFVAVMGDVGHLMLFLALAM 420

QY 421 VLAENRPAVKAQNEIWTQTFRGRYLLMLGLPSIYTGFIYNECFSRATSIFFSGWSVAA 480

DB 421 VLAENRPAVKAQNEIWTQTFRGRYLLMLGLPSIYTGFIYNECFSRATSIFFSGWSVAA 480

QY 481 MANQSGWSDAFLAQHTMTLDPNVTGVFLGPYPFGIDPIWISLAANHLSFLNSFKMKMSVI 540

DB 481 MANQSGWSDAFLAQHTMTLDPNVTGVFLGPYPFGIDPIWISLAANHLSFLNSFKMKMSVI 540

QY 541 LGVVMHAFGVVLGVFNHVFQGRHRLLETLPELTFLGLFGYLVFLVIYKWLVCWAAARA 600

DB 541 LGVVMHAFGVVLGVFNHVFQGRHRLLETLPELTFLGLFGYLVFLVIYKWLVCWAAARA 600

QY 601 ASPSILIHFINMFLFSHPNRLLYPROEVVQATLVVLALAMPVILLGLTPLHLHRRR 660

DB 601 ASPSILIHFINMFLFSHPNRLLYPROEVVQATLVVLALAMPVILLGLTPLHLHRRR 660

QY 661 RLRRRPAQREKAGLLDLPDASVNGWSDEEKAGGLDDEEAELVPSEVLHQAIHTI 720

DB 661 RLRRRPAQREKAGLLDLPDASVNGWSDEEKAGGLDDEEAELVPSEVLHQAIHTI 720

QY 721 EFCGCVSNTASYRLRWALSLAHQAQSEVLWAMVMRIGLGLGREGVGAADVLPPIFAAFA 780

DB 721 EFCGCVSNTASYRLRWALSLAHQAQSEVLWAMVMRIGLGLGREGVGAADVLPPIFAAFA 780

QY 781 VMTVAILLVMEGLSAFLHRLHWHVEFQNKFSYGTGYKLSPP 822

DB 781 VMTVAILLVMEGLSAFLHRLHWHVEFQNKFSYGTGYKLSPP 822

Db 781 VMTVAILLVMEGLSAFLHALRLHWVEFQNKFYSGTGKLSPP 822

RESULT 4

ADY92606 standard; protein; 822 AA.

XX AC ADY92606;

XX 02-JUN-2005 (first entry)

XX DE Human 116 kD osteoclast proton pump OC-116 subunit protein.

XX XX cloning; osteopathic; antiarthritic; gene therapy;

KW osteoclast proton pump; bone disease; osteoporosis; osteoarthritis.

XX XX Homo sapiens.

XX XX US2005064448-A1.

XX XX 24-MAR-2005.

XX XX 20-FEB-2004; 2004US-00783519.

XX XX 22-FEB-1996; 96US-00605378.

XX XX 18-JUL-2000; 2000US-00618304.

XX XX (STAS/) STASHENKO P.

XX XX (LIYY/) LI Y.

XX XX Stashenko P, Li Y;

XX XX WPI; 2005-241271/25.

XX XX N-PSDB; ADY92605.

XX XX New isolated genes and encoded human 116-kDa osteoclast proton pump
PT subunit polypeptides useful for treating bone mass disorders associated
PT with aberrant rate of bone degradation (e.g. osteoporosis or
PT osteoarthritis).

XX XX Claim 18; SEQ ID NO 2; 17pp; English.

XX XX The invention relates to an isolated gene encoding a polypeptide which is
CC a human 116-kD osteoclast proton pump subunit. The composition and
CC methods are useful for treating bone mass disorders characterized by
CC aberrant rate of bone degradation (e.g. osteoporosis or osteoarthritis).
CC The protein may also be used as cell surface markers for osteoclasts.
CC This sequence corresponds to the novel OC-116 proton pump subunit. The
CC encoding cDNA was isolated by differential screening of a human
CC osteoclastoma cDNA library.

XX XX Sequence 822 AA;

Query Match 100.0%; Score 822; DB 9; Length 822;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGSFRSEVALVQLPPTAAATCVSRIGELVFEFDLNASVSARFRFVVDVWRCCE 60

Db 1 MGSFRSEVALVQLPPTAAATCVSRIGELVFEFDLNASVSARFRFVVDVWRCCE 60

QY 61 LEKTTFLQEEVRRAGLVLPKGRLPAPPDRLRLIQETELRLAQELRDVRGNOQLRA 120

Db 61 LEKTTFLQEEVRRAGLVLPKGRLPAPPDRLRLIQETELRLAQELRDVRGNOQLRA 120

QY 121 QLHQQLHAAVLRQGHPEPOLAAHTDGASERTPLQAPGGPHQDLRVNFVAGVBPBKAP 180

Db 121 QLHQQLHAAVLRQGHPEPOLAAHTDGASERTPLQAPGGPHQDLRVNFVAGVBPBKAP 180

QY 181 ALERLLWRACRGLTASPRELQPLEHPVTGTPATWMTFLISYWGEOIGQKIRKTDGCF 240

Db 181 ALERLLWRACRGLTASPRELQPLEHPVTGTPATWMTFLISYWGEOIGQKIRKTDGCF 240

QY 241 CHVFPFLQOEERLGLAQLOOQSOBLOEVLGETERFLSQVLGRVLQLLPQGQVQVHKMK 300
Db 241 CHVFPFLQOEERLGLAQLOOQSOBLOEVLGETERFLSQVLGRVLQLLPQGQVQVHKMK 300
QY 301 AVYLALNQCSTVTHKCLIAEAWCSVRDLPALQEAALRDSMEEGVSAAHRIPCRDMPT 360
Db 301 AVYLALNQCSTVTHKCLIAEAWCSVRDLPALQEAALRDSMEEGVSAAHRIPCRDMPT 360
QY 361 LIRTNRTASFOGIVDRYGVGRYQEVNPAPYTIITFPFLFAVMFGDVGHLMLFLALAM 420
Db 361 LIRTNRTASFOGIVDRYGVGRYQEVNPAPYTIITFPFLFAVMFGDVGHLMLFLALAM 420
QY 421 VLAENRPAVKAQNEIWTQFFRGRYLLLMGLFSIYTGFIYNECFSRATSIIPSGWSVAA 480
Db 421 VLAENRPAVKAQNEIWTQFFRGRYLLLMGLFSIYTGFIYNECFSRATSIIPSGWSVAA 480
QY 481 MANQSGWSDAFLAQHTMLTLDPNVTGVFLGPPYFGDIPDIWLSAANHLSPFSGKMSVI 540
Db 481 MANQSGWSDAFLAQHTMLTLDPNVTGVFLGPPYFGDIPDIWLSAANHLSPFSGKMSVI 540
QY 541 LGVVMHAFGVILGVFNHVFQGRHRLILLETLPETFLGLFGYLVFLVIYKMLCVMAARA 600
Db 541 LGVVMHAFGVILGVFNHVFQGRHRLILLETLPETFLGLFGYLVFLVIYKMLCVMAARA 600
QY 601 ASPSILIHFINMFLFSHSPSNRLLYPROEVQATLVVLALAMVPILLGTPLHLHRHR 660
Db 601 ASPSILIHFINMFLFSHSPSNRLLYPROEVQATLVVLALAMVPILLGTPLHLHRHR 660
QY 661 RLRRRPAQRQENKAGLLDLPDASVNGWSSDEKAGGLDDEEAEELVPSEVLHQAIHTI 720
Db 661 RLRRRPAQRQENKAGLLDLPDASVNGWSSDEKAGGLDDEEAEELVPSEVLHQAIHTI 720
QY 721 EFCLGCVSNTASYLRMLWLSLAHQSLSEVLWAMVMRIGLGLGREGVGAANVLVPIFAA 780
Db 721 EFCLGCVSNTASYLRMLWLSLAHQSLSEVLWAMVMRIGLGLGREGVGAANVLVPIFAA 780
QY 781 VMTVAILLVMEGLSAFLHALRLHWVEFQNKFYSGTGKLSPP 822
Db 781 VMTVAILLVMEGLSAFLHALRLHWVEFQNKFYSGTGKLSPP 822

RESULT 5

ADY92606

ID ADI28834 standard; protein; 830 AA.

XX AC ADI28834;

XX DT 22-APR-2004 (first entry)

XX XX Human modifier of p53 (MP53) TCIRG-1.

XX KW Human; MP53; modifier of p53; p53; cytostatic; gene therapy;

XX KW T-cell immune regulator 1; TCIRG-1.

XX OS Homo sapiens.

XX XX WO2004004766-A1.

XX PD 15-JAN-2004.

XX PF 09-JUL-2003; 2003WO-US021378.

XX PR 10-JUL-2002; 2002US-0394992P.

XX PR 07-AUG-2002; 2002US-0401604P.

XX PR 16-SEP-2002; 2002US-0410988P.

XX PR 25-NOV-2002; 2002US-0428837P.

XX PA (EXEL-) EXELIXIS INC.

XX PI Costa MA, Maxwell ME, Lackner MR, Hung T, O'brien CL, Jin Y;

XX PI Nicoll M, Hai B, Zhang H, Lickteig K, Amundeen CD;

XX XX WPI; 2004-142922/14.

DR N-PSDB; ADI28820.
 DR GENBANK; 19924145.
 XX
 PT Identifying a candidate p53 pathway modulating agent for treating e.g.,
 PT cancer, comprises contacting an assay system comprising a MP53
 PT polypeptide or nucleic acid with a test agent and detecting a test agent-
 PT biased activity.
 XX
 XX Example 2; SEQ ID NO 25; 139pp; English.
 XX
 CC In the present invention, genetic screens were designed to identify
 CC modifiers of the p53 pathway in caenorhabditis elegans, where a
 CC homozygous p53 deletion mutant was used. Various specific genes were
 CC silenced by RNA inhibition. Genes causing altered phenotypes in the worms
 CC were identified as modifiers of the p53 pathway. Human orthologs, denoted
 CC modifiers of p52 (MP53), of these modifiers were identified. These
 CC include the present sequence, characterised as T-cell immune regulator 1,
 CC ATPase, H+ transporting, or lysosomal V0 protein a isoform 3. MP53
 CC nucleic acids and polypeptides are attractive drug targets for the
 CC treatment of pathologies associated with a defective p53 signalling
 CC pathway, such as cancer. Methods for modulating MP53-function and/or the
 CC p53 pathway in a mammalian cell involve contacting the cell with an agent
 CC that specifically binds a MP53 polypeptide or nucleic acid. The agent may
 CC be a small molecular modulator, a nucleic acid modulator or an antibody.
 XX
 XX Sequence 830 AA;
 SQ
 Query Match 66.3%; Score 545; DB 8; Length 830;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 765; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 57 RCELEKTFPQEEVRRAGLVLPKGRLPAPPPDRDLRIQETETERLAQELDRVGNQ 116
 57 RCELEKTFPQEEVRRAGLVLPKGRLPAPPPDRDLRIQETETERLAQELDRVGNQ 116
 117 ALPAQLHQLHAQVLRQGHPEPOLAAHTDGAERTLLQAPCGPHQDLRVNFVAGV 176
 117 ALPAQLHQLHAQVLRQGHPEPOLAAHTDGAERTLLQAPCGPHQDLRVNFVAGV 176
 177 HKPALERLLWRACRGLFIASFRELEQPLEHPVTGEPATWTFILSYGQIQKIRKIT 236
 177 HKPALERLLWRACRGLFIASFRELEQPLEHPVTGEPATWTFILSYGQIQKIRKIT 236
 237 DCFHCHVFPPLQEEARLQALQOQOSELQVGLGTERFLSQVLRVLLPQGVQV 296
 237 DCFHCHVFPPLQEEARLQALQOQOSELQVGLGTERFLSQVLRVLLPQGVQV 296
 297 HKMKAVYALNOCVSTTHKCLIAEAWCSVRDLPAQELARDSSMEEGVSAVAHRIPCRD 356
 297 HKMKAVYALNOCVSTTHKCLIAEAWCSVRDLPAQELARDSSMEEGVSAVAHRIPCRD 356
 357 MPPTLIRTNRTASFGQIVDRYGVGRYQEVNRPAPYTIITPFLFVAVMFGDVGHGLMFLP 416
 357 MPPTLIRTNRTASFGQIVDRYGVGRYQEVNRPAPYTIITPFLFVAVMFGDVGHGLMFLP 416
 417 ALAMVLAENRPAKAAQNEIWTQFFRGYRLLMLGLFSIYGTIYNECFSRATSIIPSGW 476
 417 ALAMVLAENRPAKAAQNEIWTQFFRGYRLLMLGLFSIYGTIYNECFSRATSIIPSGW 476
 477 SVAAANQSGWSDAFLAQHTMLTLDNVTGVFLGPPYGGIDPIWLSLAANHLSTLNGFKMK 536
 477 SVAAANQSGWSDAFLAQHTMLTLDNVTGVFLGPPYGGIDPIWLSLAANHLSTLNGFKMK 536
 537 MSVILGVVHMAFGVGLGVFNHVFQGRHRLLETLPELTFLGLFGLVFLVYKMLCVW 596
 537 MSVILGVVHMAFGVGLGVFNHVFQGRHRLLETLPELTFLGLFGLVFLVYKMLCVW 596
 597 AARAAS-PSILIHFINFLSPSNRLLYPRQEVQATLVIALAMVPIILLGTPLHLL 655
 597 AARAASPSILIHFINFLSPSNRLLYPRQEVQATLVIALAMVPIILLGTPLHLL 656
 656 HRHRRRLRRPADRQENKAGLLDLPASVNGWSSDEEKAGGLDDEEAEELVSEVLMHQ 715

Db 657 HRHRRRLRRPADRQENKAGLLDLPASVNGWSSDEEKAGGLDDEEAEELVSEVLMHQ 716
 QY 716 AIHTIFCLGCVNTASYLRLWALSALAHQSEVLWAMVMRIGLGLGREVGVAAVLVPI 775
 Db 717 AIHTIFCLGCVNTASYLRLWALSALAHQSEVLWAMVMRIGLGLGREVGVAAVLVPI 776
 QY 776 FAAPAVMTVAILLVMEGLSAPFLHALRLHWVEFQNKFYSGTGKLSPP 822
 Db 777 FAAPAVMTVAILLVMEGLSAPFLHALRLHWVEFQNKFYSGTGKLSPP 823
 RESULT 6
 ADPI2465
 ID ADPI2465 standard; protein; 830 AA.
 XX
 AC ADPI2465;
 DT 12-AUG-2004 (first entry)
 XX
 DE Protein encoded by mRNA of the invention #75.
 XX
 KW transplant rejection; immune system; rheumatoid arthritis; lupus;
 KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS.
 XX
 OS Homo sapiens.
 XX
 PN WO2004042346-A2.
 XX
 PD 21-MAY-2004.
 XX
 PF 24-APR-2003; 2003WO-US012946.
 XX
 PR 24-APR-2002; 2002US-00131831.
 PR 20-DEC-2002; 2002US-00325899.
 XX
 PA (EXPR-) EXPRESSION DIAGNOSTICS INC.
 XX
 PI Wohlgenuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
 PI Rosenberg S;
 XX
 XX WPI; 2004-400724/37.
 XX
 PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
 PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
 PT rejection, in an individual, comprises detecting the expression level of
 PT the genes.
 XX
 PS Claim 65; SEQ ID NO 2474; 1762pp; English.
 XX
 CC The present invention relates to diagnosing or monitoring transplant
 CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
 CC comprising detecting the expression level of one or more genes. The
 CC methods, system and kits are useful in diagnosing or monitoring
 CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
 CC islet, lung, bone marrow or stem cell transplant rejection.
 CC xenotransplant rejection or mechanical organ replacement rejection, in an
 CC individual. The method is also useful in assessing the immune status of
 CC an individual. The methods are also useful in diagnosing and monitoring
 CC diseases that involve the immune system, e.g. rheumatoid arthritis,
 CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or
 CC viral, bacterial or fungal infection. The present sequence represents a
 CC protein that is encoded by the mRNA of the invention.
 XX
 SQ Sequence 830 AA;

Query Match 66.3%; Score 545; DB 8; Length 830;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 765; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 57 RCELEKTFPQEEVRRAGLVLPKGRLPAPPPDRDLRIQETETERLAQELDRVGNQ 116
 Db 57 RCELEKTFPQEEVRRAGLVLPKGRLPAPPPDRDLRIQETETERLAQELDRVGNQ 116

```
QY 117 ALRAQLHQLHAAVLROGHEPOLAAHTDASERTPLLAQPGPHQDLRVNMFVAGAVEP 176
Db 117 ALRAQLHQLHAAVLROGHEPOLAAHTDASERTPLLAQPGPHQDLRVNMFVAGAVEP 176
QY 177 HKPALERLLWRACRGFLIASPRELEQPLEHPVTGEPATWMTFLISYWGEGIQGKIRKIT 236
Db 177 HKPALERLLWRACRGFLIASPRELEQPLEHPVTGEPATWMTFLISYWGEGIQGKIRKIT 236
QY 237 DCFHCHVFPFLOQEEARLGAALQOQSQEQLQEVLTGETERFSLQVLRVQLLPPQGVQV 296
Db 237 DCFHCHVFPFLOQEEARLGAALQOQSQEQLQEVLTGETERFSLQVLRVQLLPPQGVQV 296
QY 297 HKMKAVYLALNQCSTVTHKCLIAEAWCSVRDLPALQEARLDRSSMEEGVSAVAHRIPCRD 356
Db 297 HKMKAVYLALNQCSTVTHKCLIAEAWCSVRDLPALQEARLDRSSMEEGVSAVAHRIPCRD 356
QY 357 MPPTLIRNRTASFGQIVDRYGVGRYQEVNPPATYITITPPFLFAMFGDVGHGLMLFLF 416
Db 357 MPPTLIRNRTASFGQIVDRYGVGRYQEVNPPATYITITPPFLFAMFGDVGHGLMLFLF 416
QY 417 ALAMVLAENRPAVKAAQNEIWQTFPRGRYLLMLGLFSIYTGFIYNECSRATSIPPSGW 476
Db 417 ALAMVLAENRPAVKAAQNEIWQTFPRGRYLLMLGLFSIYTGFIYNECSRATSIPPSGW 476
QY 477 SVAAMNQGSGWDAFLAQTMLTLDPNVTGVFLGYPFGIDPTWLSLAANHLSFLNSFKMK 536
Db 477 SVAAMNQGSGWDAFLAQTMLTLDPNVTGVFLGYPFGIDPTWLSLAANHLSFLNSFKMK 536
QY 537 MSVILGVVHMAQVVLGVNHNHFGQRHLLLETLPFLGLGLFGLYVFLVIYKWLQVW 596
Db 537 MSVILGVVHMAQVVLGVNHNHFGQRHLLLETLPFLGLGLFGLYVFLVIYKWLQVW 596
QY 597 AARAS-PSILTHFIMFLPSHSPNRLLYPROEVVQATLVLAAMVPILLLGTPHL 655
Db 597 AARAS-PSILTHFIMFLPSHSPNRLLYPROEVVQATLVLAAMVPILLLGTPHL 655
QY 656 HRRHRRRRPADROENKAGLLDLPDASVNGSSDEEKAGGLDDEEAEALVPSEVLMHQ 715
Db 657 HRRHRRRRPADROENKAGLLDLPDASVNGSSDEEKAGGLDDEEAEALVPSEVLMHQ 716
QY 716 ALHTTEFCGCVSNTASYLRLLWALSIAHAQLSEVLWAMVMRIGLGLGREVGVAALVPI 775
Db 717 ALHTTEFCGCVSNTASYLRLLWALSIAHAQLSEVLWAMVMRIGLGLGREVGVAALVPI 776
QY 776 PAAPAVMTVAILLVMEGLSAPFLHALRLHWWVEFQNKFGYSGYKLSPP 822
Db 777 PAAPAVMTVAILLVMEGLSAPFLHALRLHWWVEFQNKFGYSGYKLSPP 823

RESULT 7
ABM81281
XX ABM81281 standard; protein; 830 AA.
AC ABM81281;
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) polypeptide PRO82023, SEQ:3305.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
```

```
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH ) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
XX
DR WPI; 2004-347921/32.
XX N-PSDB; ACN39267.
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
XX Claim 12; SEQ ID NO 3305; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention
XX
XX Sequence 830 AA;
XX
Query Match 66.3%; Score 545; DB 8; Length 830;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 765; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 57 RCELEKTFTLQEEVRRAGLVLPKPKRLPAPPDRDLRLRQETRLAQELRDVRGNQ 116
Db 57 RCELEKTFTLQEEVRRAGLVLPKPKRLPAPPDRDLRLRQETRLAQELRDVRGNQ 116
QY 117 ALRAQLHQLHAAVLROGHEPOLAAHTDASERTPLLAQPGPHQDLRVNMFVAGAVEP 176
Db 117 ALRAQLHQLHAAVLROGHEPOLAAHTDASERTPLLAQPGPHQDLRVNMFVAGAVEP 176
QY 177 HKPALERLLWRACRGFLIASPRELEQPLEHPVTGEPATWMTFLISYWGEGIQGKIRKIT 236
Db 177 HKPALERLLWRACRGFLIASPRELEQPLEHPVTGEPATWMTFLISYWGEGIQGKIRKIT 236
QY 237 DCFHCHVFPFLOQEEARLGAALQOQSQEQLQEVLTGETERFSLQVLRVQLLPPQGVQV 296
Db 237 DCFHCHVFPFLOQEEARLGAALQOQSQEQLQEVLTGETERFSLQVLRVQLLPPQGVQV 296
QY 297 HKMKAVYLALNQCSTVTHKCLIAEAWCSVRDLPALQEARLDRSSMEEGVSAVAHRIPCRD 356
Db 297 HKMKAVYLALNQCSTVTHKCLIAEAWCSVRDLPALQEARLDRSSMEEGVSAVAHRIPCRD 356
QY 357 MPPTLIRNRTASFGQIVDRYGVGRYQEVNPPATYITITPPFLFAMFGDVGHGLMLFLF 416
Db 357 MPPTLIRNRTASFGQIVDRYGVGRYQEVNPPATYITITPPFLFAMFGDVGHGLMLFLF 416
QY 417 ALAMVLAENRPAVKAAQNEIWQTFPRGRYLLMLGLFSIYTGFIYNECSRATSIPPSGW 476
Db 417 ALAMVLAENRPAVKAAQNEIWQTFPRGRYLLMLGLFSIYTGFIYNECSRATSIPPSGW 476
```

QY 477 SVAAANQSGSDAFLAQTMLTLDPNVTGVFLGYPFGIDPIWSLAANHLSPLNSPKMK 536
DB 477 SVAAANQSGSDAFLAQTMLTLDPNVTGVFLGYPFGIDPIWSLAANHLSPLNSPKMK 536
QY 537 MSVILGVVHMAFGVVLGVFNHVFQGRHRLILLETLPETLFLGLFGVLVFLVIYKWLVCW 596
DB 537 MSVILGVVHMAFGVVLGVFNHVFQGRHRLILLETLPETLFLGLFGVLVFLVIYKWLVCW 596
QY 597 AARAAS-PSILIHFINFLSHSPSNRLLYPRQEVQATLVVLALAMVPILLGTPLHL 655
DB 597 AARAASPSILIHFINFLSHSPSNRLLYPRQEVQATLVVLALAMVPILLGTPLHL 656
QY 656 HRHRRRLRRPADQENKAGLLDLPASVNGSSDEEKAGGLDDEEAELVPSEVLMLHQ 715
DB 657 HRHRRRLRRPADQENKAGLLDLPASVNGSSDEEKAGGLDDEEAELVPSEVLMLHQ 716
QY 716 AIHTIEFCGCVSNTASYLRILWALSALAHQSEVLWAMVMRIGLIGRGVGVAAVVLVPI 775
DB 717 AIHTIEFCGCVSNTASYLRILWALSALAHQSEVLWAMVMRIGLIGRGVGVAAVVLVPI 776
QY 776 FAFAVMTVAILLVMEGLSAPFLHALRLHWHVEFQNKFSYGTGYKLSPP 822
DB 777 FAFAVMTVAILLVMEGLSAPFLHALRLHWHVEFQNKFSYGTGYKLSPP 823

RESULT 8

ADH22516
ID ADH22516 standard; protein; 787 AA.
XX
AC ADH22516;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human transporter & ion channel (TRICH) protein SeqID14.
XX
KW human; transporters and ion channel; TRICH; cell proliferative;
KW arteriosclerosis; cancer; autoimmune/inflammatory; AIDS; asthma;
KW neurological; epilepsy; stroke; developmental; Cushing's syndrome;
KW hypothyroidism; infection; gene therapy; cytostatic; antiinflammatory;
KW immunosuppressive; antiaesthetic; anticonvulsant; nootropic;
KW neuroprotective; single nucleotide polymorphism; SNP.
XX
OS Homo sapiens.
XX
PN WO2003093444-A2.
XX
PD 13-NOV-2003.
XX
PF 02-MAY-2003; 2003WO-US014026.
XX
PR 03-MAY-2002; 2002US-0377435P.
PR 03-MAY-2002; 2002US-0377444P.
PR 05-JUN-2002; 2002US-0386497P.
PR 11-JUN-2002; 2002US-0388180P.
XX
XX (INCY-) INCYTE CORP.
XX
XX Baughn MR, Becha SD, Bulloch SA, Chang H, Elliott VS;
PI Emerling BM, Griffin JA, Hafalia AJA, Ison CH, Jackson AA, Jiang X;
PI Jin P, Kable AE, Khare R, Lee SY, Lee S, Mason PM, Marquis JP;
PI Ramkumar J, Richardson TW, Swarnakar A, Tran UK, Chawla NK;
PI Wilson AD;
XX
XX WPI; 2004-022655/02.
DR N-PSDB; ADH22582.
XX
XX New human transporters and ion channels (TRICH), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant TRICH expression e.g. cancer, AIDS, arteriosclerosis, epilepsy,
PT or infections.
XX
XX Claim 1; SEQ ID NO 14; 448pp; English.

CC This invention relates to novel isolated polynucleotides identified as
CC human transporters and ion channels (TRICH), and the encoded polypeptides
CC thereof. Specifically, it describes using these TRICH molecules, as well
CC as agonists, antagonists, antibodies, expression vectors and host cells,
CC in appropriate screening and toxicity assays to assess the effects of
CC exogenous compounds on TRICH expression. The present invention describes
CC TRICH compositions that are useful in the diagnosis, treatment and
CC prevention of various disorders such as cell proliferative (e.g.
CC arteriosclerosis, cancer), autoimmune/inflammatory (e.g. AIDS, asthma),
CC neurological (e.g. epilepsy, stroke) and developmental (e.g. Cushing's
CC syndrome hypothyroidism) and for infections. Accordingly, these TRICH
CC molecules can be used for gene therapy purposes and exhibit various
CC activities such as cytostatic, antiinflammatory, immunosuppressive,
CC antiaesthetic, anticonvulsant, nootropic and neuroprotective.
CC Furthermore, a microarray is useful in monitoring or measuring protein-
CC protein interactions, drug-target interactions and gene expression
CC profiles. This polypeptide sequence is a human TRICH protein of the
CC invention.
XX

SQ Sequence 787 AA;

Query Match 60.9%; Score 501; DB 8; Length 787;
Best Local Similarity 99.8%; Pred No. 0;
Matches 601; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MGSMPRSEEVAVQLFLPTAAAYTCVSRGELGLVEFRDLNASVSARFRRVVDVWRCCE 60
DB 1 MGSMPRSEEVAVQLFLPTAAAYTCVSRGELGLVEFRDLNASVSARFRRVVDVWRCCE 60
QY 61 LEKTFTEQBEVVRAGLVLPKGRLPAPPRDLRIQETTERLAQLRDRVGNQOALRA 120
DB 61 LEKTFTEQBEVVRAGLVLPKGRLPAPPRDLRIQETTERLAQLRDRVGNQOALRA 120
QY 121 QLHQLQHLAAVLRQGHPEPQLAAATDCASERTPLQAPGGPHQDLRVNFVAGAVEPHKAP 180
DB 121 QLHQLQHLAAVLRQGHPEPQLAAATDCASERTPLQAPGGPHQDLRVNFVAGAVEPHKAP 180
QY 181 ALBRLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEOIGQKIRKITDCFH 240
DB 181 ALBRLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEOIGQKIRKITDCFH 240
QY 241 CHVFPFLOQBEARLQALQLOQSOELOEVLETGERFSLQVLGRVLLQPPGVQVHKMK 300
DB 241 CHVFPFLOQBEARLQALQLOQSOELOEVLETGERFSLQVLGRVLLQPPGVQVHKMK 300
QY 301 AVYALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLDSMEEGVSAVAHRIPCRDMPPT 360
DB 301 AVYALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLDSMEEGVSAVAHRIPCRDMPPT 360
QY 361 LIRTNRTASFOGIVDRYGVGRYQEVNPAPTYITTPPFLFAVMFGDVGHGLLMFLFALAM 420
DB 361 LIRTNRTASFOGIVDRYGVGRYQEVNPAPTYITTPPFLFAVMFGDVGHGLLMFLFALAM 420
QY 421 VLAENRPAVKAAQNEIWQTFRRGRYLLMLLSIYTGFIYNECFSRATSIFFSGWSVAA 480
DB 421 VLAENRPAVKAAQNEIWQTFRRGRYLLMLLSIYTGFIYNECFSRATSIFFSGWSVAA 480
QY 481 MANQSGWSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWSLAANHLSFLNSPKMKSVI 540
DB 481 MANQSGWSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWSLAANHLSFLNSPKMKSVI 540
QY 541 LGVHVHMAFGVVLGVFNHVFQGRHRLILLETLPETLFLGLFGVLVFLVIYKWLVCWAARA 600
DB 541 LGVHVHMAFGVVLGVFNHVFQGRHRLILLETLPETLFLGLFGVLVFLVIYKWLVCWAARA 600
QY 601 AS 602
DB 601 AS 602

RESULT 9

ABM84820
ID ABM84820 standard; protein; 787 AA.

XX AC ABM84820;
XX DT 18-NOV-2004 (first entry)
XX DE Human diagnostic and therapeutic pprotein SEQ ID NO:5069.
XX KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX OS Homo sapiens.
XX PN WO2004023973-A2.
XX PD 25-MAR-2004.
XX PF 12-SEP-2003; 2003WO-US0282227.
XX PR 12-SEP-2002; 2002US-0410259P.
XX PR 12-SEP-2002; 2002US-0410260P.
XX PA (INCY-) INCYTE CORP.
XX PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthorne TA, Suchorski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patry S, Shi X, Suarez CJ;
XX WPI; 2004-329368/30.
XX DR N-PSDB; ACN43472.
XX PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
XX in diagnosing a condition, disease or disorder associated with human
XX molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
XX in gene mapping.
XX PS Claim 27; Page; 190pp; English.
XX CC The invention relates to novel diagnostic and therapeutic polynucleotides
XX selected from one of the 2722 sequences defined in the specification. A
XX polynucleotide of the invention may have a use in gene therapy. The human
XX diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
XX used to diagnose a particular condition, disease or disorder associated
XX with human molecules, e.g. cell proliferative disorders,
XX autoimmune/inflammatory disorder, developmental disorder, endocrine
XX disorder, neurological disorders, gastrointestinal disorders, or
XX infections caused by virus, bacteria, fungi or parasite. The dithp
XX molecules may also be used in genetic mapping, in identifying individuals
XX from minute biological samples, in detecting single nucleotide
XX polymorphisms, as molecular weight markers, and for somatic or germline
XX gene therapy. The present sequence represents a dithp protein of the
XX invention. Note: The sequence data for this patent is not represented in
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX SQ Sequence 787 AA;
Query Match 54.1%; Score 445; DB 8; Length 787;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 545; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 57 RCELEKFTTFLQEEVRRAGLVLPKPKGLPAPPPRDLRLRIQETTERLAQELRDVRGNOQ 116
DB 57 RCELEKFTTFLQEEVRRAGLVLPKPKGLPAPPPRDLRLRIQETTERLAQELRDVRGNOQ 116
QY 117 ALRAQLHQLHAAVLROGHEPQLAAAHDTGASERTLLQAPGGPHQDLRVNFVAGAVEP 176
DB 117 ALRAQLHQLHAAVLROGHEPQLAAAHDTGASERTLLQAPGGPHQDLRVNFVAGAVEP 176
QY 177 HKPALERLLMRACRGLFIASPRELEQPLEHPVTGEPATWMTFLISYWGEGIQGKIRKIT 236

DB 177 HKPALERLLMRACRGLFIASPRELEQPLEHPVTGEPATWMTFLISYWGEGIQGKIRKIT 236
QY 237 DCFHCHVFPFLQEEARLQALQLOOQSOQELQEVLTETFLSQVLGRVLQQLPPGQVQV 296
DB 237 DCFHCHVFPFLQEEARLQALQLOOQSOQELQEVLTETFLSQVLGRVLQQLPPGQVQV 296
QY 297 HOKKAVYLALNQCSTVTHKCLIAEAWCSVRDLPALQEBALRDSSMBEGVSAVAHRIPCRD 356
DB 297 HOKKAVYLALNQCSTVTHKCLIAEAWCSVRDLPALQEBALRDSSMBEGVSAVAHRIPCRD 356
QY 357 MPPTLIRTNRTASFGQIVDRYGVGRYQVBNPAPVTIITFPFLFAVMFGDVGHLLMELF 416
DB 357 MPPTLIRTNRTASFGQIVDAVGVRYQVBNPAPVTIITFPFLFAVMFGDVGHLLMELF 416
QY 417 ALAMVLAENRPAVKAQAONEIMQTFPRGRYLLLMGLFSIYTGFIYNECFSRATSIFFSGW 476
DB 417 ALAMVLAENRPAVKAQAONEIMQTFPRGRYLLLMGLFSIYTGFIYNECFSRATSIFFSGW 476
QY 477 SVAAMANQSGWSDAFLAQTMTLTDPNVTGVFLGYPFGIDDIWSLAANHLSFLNSFKMK 536
DB 477 SVAAMANQSGWSDAFLAQTMTLTDPNVTGVFLGYPFGIDDIWSLAANHLSFLNSFKMK 536
QY 537 MSVILGVVHMAFGVVLGVFNHVFQGRHLLLETPELTFLGLFGYLVFLVIYKWLQVW 596
DB 537 MSVILGVVHMAFGVVLGVFNHVFQGRHLLLETPELTFLGLFGYLVFLVIYKWLQVW 596
QY 597 AARAAS 602
DB 597 AARAAS 602
RESULT 10
AAW98112
ID AAW98112 standard; protein; 614 AA.
XX AC AAW98112;
XX DT 21-JUN-1999 (first entry)
XX DE T-cell membrane protein TIRC7.
XX KW TIRC7; T-cell immune response cDNA 7; T-cell membrane protein; human;
KW immunosuppressive; organ transplantation; graft rejection;
KW transplant rejection; autoimmune disease; allergy; infection; tumour;
KW cancer; rheumatoid arthritis; lupus erythematosus; multiple sclerosis;
KW encephalitis; vasculitis; diabetes mellitus; pancreatitis; gastritis;
KW thyroiditis; asthma; lepraematosis; therapy.
XX OS Homo sapiens.
XX FH Location/Qualifiers
XX Protein 1..601
XX /note= "separately claimed in Claim 1"
XX Domain 173..209
XX /note= "transmembrane domain"
XX Domain 229..246
XX /note= "transmembrane domain"
XX Domain 319..346
XX /note= "transmembrane domain"
XX Domain 356..400
XX /note= "transmembrane domain"
XX Domain 413..437
XX /note= "transmembrane domain"
XX Domain 513..532
XX /note= "transmembrane domain"
XX Domain 544..585
XX /note= "transmembrane domain"
XX WO9911782-A1.
XX 11-MAR-1999.

PF 28-AUG-1998; 98WO-EF005462.
XX
PR 29-AUG-1997; 97DE-01038710.
PR 12-FEB-1998; 98DE-02002653.
XX
PA (BGHM) BRIGHAM & WOMENS HOSPITAL.
PA (UTKU/) UTKU N.
XX
XX Utku N, Gullans SR, Milford EL;
PI WPI; 1999-205186/17.
XX N-PSDB; AAX24912.
DR
DR A novel T-cell transmembrane protein (TIRC7) - useful for modulating
PT immune responses, for use in organ transplantations and treatment of
PT infectious disease.
XX
XX Claim 1(i); Page 84-86; 97pp; English.
XX
XX The present sequence is TIRC7, a novel T-cell transmembrane protein that
CC exhibits a central role in T-cell activation in vitro and in vivo. TIRC7
CC mRNA is transiently up-regulated in the early phase of T-cell activation
CC via a calcineurin-dependent pathway. TIRC7 cDNA (see AAX24912) was
CC identified following a screen for genes that are differentially expressed
CC in alloactivated human T cells. A second TIRC7 protein (see AAW98113) is
CC identical except for one amino acid substitution (Arg121 to Gln),
CC suggesting 2 alleles of the TIRC7 gene. The invention provides vectors,
CC host cells, antibodies, TIRC7 peptides, polypeptides and polynucleotides
CC that are useful in diagnostic compositions for identifying T-cell
CC activating or co-stimulating compounds, or for identifying inhibitors of
CC T-cell activation and stimulation. TIRC7 peptides or polypeptides,
CC vectors and antibodies are useful in the preparation of pharmaceutical
CC compositions for treatment of acute and chronic diseases involving T-cell
CC activation and Th1 and Th2 immune response, for the treatment of acute
CC and chronic rejection of allo- and xeno organ transplants and bone marrow
CC transplantation, for the treatment of rheumatoid arthritis, lupus
CC erythematosus, multiple sclerosis, encephalitis, vasculitis, diabetes
CC mellitus, pancreatitis, gastritis, thyroiditis, for the treatment of
CC malign disorders of T, B or NK cells, for the treatment of asthma,
CC lepramatosi, Helicobacter pylori associated gastritis, or for the
XX treatment of skin, adrenal or lung tumors (claimed)
XX
XX Sequence 614 AA;
SQ

Query Match 46.8%; Score 385; DB 2; Length 614;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 605; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
217 MTFLLSYWGEIQGKIRKIDCFCHVFPFLOQEEARLQALQOQSQBLQVLGETER 276
1 MTFLLSYWGEIQGKIRKIDCFCHVFPFLOQEEARLQALQOQSQBLQVLGETER 60
277 FLSQVLGRVLQPLPGQVQVHKMAYVLANQCSVSTTHKCLIAEAWCSVRDLPALQAL 336
61 FLSQVLGRVLQPLPGQVQVHKMAYVLANQCSVSTTHKCLIAEAWCSVRDLPALQAL 120
337 RDSMEEGSAVAHRIPCRDMPPTLIRNTRFTASFGQIVDRYGVGRYQVNPAPYTIITP 396
121 RDSMEEGSAVAHRIPCRDMPPTLIRNTRFTASFGQIVDRYGVGRYQVNPAPYTIITP 180
397 PFLFVAFVFGDVGHGLMLFALAMVLAENRPAVKAQNEIWTQFFRGYRLLLMGLFSIY 456
181 PFLFVAFVFGDVGHGLMLFALAMVLAENRPAVKAQNEIWTQFFRGYRLLLMGLFSIY 240
457 TGIYNECSRATSIIPSGWSVAAANQSGWSDAFLAQHTMLTLDNVTGVFLGYPYPGI 516
241 TGIYNECSRATSIIPSGWSVAAANQSGWSDAFLAQHTMLTLDNVTGVFLGYPYPGI 300
517 DPTWSLAANHLFLNSFKMWSVILGVVHMAFGVLFVFNHFGQRHRLLLTLP 576
301 DPTWSLAANHLFLNSFKMWSVILGVVHMAFGVLFVFNHFGQRHRLLLTLP 360
577 LLGLFGYLVFLVYKWLVCVWAARAAS-PSILIHFINNLFSPSNRLLYPROEVVQATL 635

Db 361 LLGLFGYLVFLVYKWLVCVWAARAASPSILIHFINNLFSPSNRLLYPROEVVQATL 420
QY 636 VVLALANVPILLIGTTPHLHLHRRRLRRPADRQENKAGLLDLPDASVNGHSSDDEKA 695
Db 421 VVLALANVPILLIGTTPHLHLHRRRLRRPADRQENKAGLLDLPDASVNGHSSDDEKA 480
QY 696 GGLDDEEAELVPSEVLMHQAIHTIEFCLGCVSNTASYLRMLWALSALAHQQLSEVLMAMVM 755
Db 481 GGLDDEEAELVPSEVLMHQAIHTIEFCLGCVSNTASYLRMLWALSALAHQQLSEVLMAMVM 540
QY 756 RIGLIGRGVGVAAVVLVPIFAFAVMTVAILLVMEGLSAFLHALRLHWVEFONKFVSGT 815
Db 541 RIGLIGRGVGVAAVVLVPIFAFAVMTVAILLVMEGLSAFLHALRLHWVEFONKFVSGT 600
QY 816 GYKLSPPF 822
Db 601 GYKLSPPF 607

RESULT 11
ABM81282
ID ABM81282 standard; protein; 614 AA.
XX
XX AC ABM81282;
XX
XX 18-NOV-2004 (first entry)
XX
XX Tumour-associated antigenic target (TAT) polypeptide PRO24831, SEQ:3307.
XX
XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic.
XX
XX Homo sapiens.
XX
XX WO2004030615-A2.
XX
XX 15-APR-2004.
XX
XX 29-SEP-2003; 2003WO-US028547.
XX
XX 02-OCT-2002; 2002US-0414971P.
XX
XX (GETH) GENENTECH INC.
XX
XX Wu TD, Zhang Z, Zhou Y;
XX
XX WPI: 2004-347921/32.
XX N-PSDB; ACN39268.
XX
XX New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
XX Claim 12; SEQ ID NO 3307; 7273pp; English.
XX
XX The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,

CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention
XX

SQ Sequence 614 AA;

Query Match 46.8%; Score 385; DB 8; Length 614;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 605; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 217 MTFLISYMGEGIQGKIRKIDCFHCHVFPFLOQEEARLQALQOQSQSDELQVLGETER 276
Db 1 MTFLISYMGEGIQGKIRKIDCFHCHVFPFLOQEEARLQALQOQSQSDELQVLGETER 60
QY 277 FLSQLGRVLQLLPQGVQVHKAVYALNQCSTTHKCLIAEAWCVRDLPALQEL 336
Db 61 FLSQLGRVLQLLPQGVQVHKAVYALNQCSTTHKCLIAEAWCVRDLPALQEL 120
QY 337 RDSMEEGVSAVAHRIPCRDMPPTLIRTNRTASFGIYDVRGVGRVQVNPAPYTIIF 396
Db 121 RDSMEEGVSAVAHRIPCRDMPPTLIRTNRTASFGIYDVRGVGRVQVNPAPYTIIF 180
QY 397 PFLFVMPGDVGHLMLFALAMVLAENRPAKAAQNEIWQTFGRYLLLLMLGFSY 456
Db 181 PFLFVMPGDVGHLMLFALAMVLAENRPAKAAQNEIWQTFGRYLLLLMLGFSY 240
QY 457 TGFYNECESRATSIFFPSGWSVAANQSGWSDAFLAQHTMLTDNVTGVLGPYPFGI 516
Db 241 TGFYNECESRATSIFFPSGWSVAANQSGWSDAFLAQHTMLTDNVTGVLGPYPFGI 300
QY 517 DPWLSAANHLFLSFKMKSIVILGVHMAFGVILGVFNHVFHGORHRLLETLPELTF 576
Db 301 DPWLSAANHLFLSFKMKSIVILGVHMAFGVILGVFNHVFHGORHRLLETLPELTF 360
QY 577 LGLFGYLVFLVIYKWCVAARAAS-PSILIHFINMFLFSHPNRLLYPRQEVVQATL 635
Db 361 LGLFGYLVFLVIYKWCVAARAAS-PSILIHFINMFLFSHPNRLLYPRQEVVQATL 420
QY 636 VVALAMVPILLGLTPHLHLHRRRLRRPRADQENKAGLLDLPDASVNGWSSDEKA 695
Db 421 VVALAMVPILLGLTPHLHLHRRRLRRPRADQENKAGLLDLPDASVNGWSSDEKA 480
QY 696 GGLDDEEAELVPSVLMHQAHTIEFCIGCVSNTASYLRLWALSIAHAQLSEVLWAMVM 755
Db 481 GGLDDEEAELVPSVLMHQAHTIEFCIGCVSNTASYLRLWALSIAHAQLSEVLWAMVM 540
QY 756 RIGLGLGREGVGAUVLVPIFAFVMTVAILLVMEGLSAFLHALRLHWVEFQNFYSGT 815
Db 541 RIGLGLGREGVGAUVLVPIFAFVMTVAILLVMEGLSAFLHALRLHWVEFQNFYSGT 600
QY 816 GYKLSPP 822
Db 601 GYKLSPP 607

RESULT 12

ID AAW98113 standard; protein; 614 AA.

XX AC AAW98113;

XX DT 21-JUN-1999 (first entry)

XX DE T-cell membrane protein TIRC7.

XX TIRC7; T-cell immune response cDNA 7; T-cell membrane protein; human;
KW immunosuppressive; organ transplantation; graft rejection.

KW transplant rejection; autoimmune disease; allergy; infection; tumour;
KW cancer; rheumatoid arthritis; lupus erythematosus; multiple sclerosis;
KW encephalitis; vasculitis; diabetes mellitus; pancreatitis; gastritis;
KW thyroiditis; asthma; lepramatosi; therapy.

OS Homo sapiens.

PH Key Location/Qualifiers
FT Protein 1. .601
FT /note= "separately claimed in Claim 1"
FT Domain 173. .209
FT /note= "transmembrane domain"
FT Domain 229. .246
FT /note= "transmembrane domain"
FT Domain 319. .346
FT /note= "transmembrane domain"
FT Domain 356. .400
FT /note= "transmembrane domain"
FT Domain 413. .437
FT /note= "transmembrane domain"
FT Domain 513. .532
FT /note= "transmembrane domain"
FT Domain 544. .585
FT /note= "transmembrane domain"

XX WO9911782-A1.

PN 11-MAR-1999.

XX 28-AUG-1998; 98WO-EP005462.

XX 29-AUG-1997; 97DE-01038710.

XX 12-FEB-1998; 98DE-02002653.

XX (BGHM) BRIGHAM & WOMENS HOSPITAL.
XX (UTKU/) UTKU N.

XX Utku N, Gullans SR, Milford EL;

XX WPI; 1999-205186/17.

XX N-PSDB; AAX24913.

XX A novel T-cell transmembrane protein (TIRC7) - useful for modulating
XX immune responses, for use in organ transplantations and treatment of
XX infectious disease.

XX Claim 1(i); Page 93-95; 97pp; English.

XX The present sequence is TIRC7, a novel T-cell transmembrane protein that
XX exhibits a central role in T-cell activation in vitro and in vivo. TIRC7
XX mRNA is transiently up-regulated in the early phase of T-cell activation
XX via a calcineurin-dependent pathway. TIRC7 cDNA (see AAX24912) was
XX identified following a screen for genes that are differentially expressed
XX in activated human T cells. A second TIRC7 protein (see AAW98112) is
XX identical except for one amino acid substitution (Gln121 to Arg).

XX suggesting 2 alleles of the TIRC7 gene. The invention provides vectors,
XX host cells, antibodies, TIRC7 peptides, polypeptides and polynucleotides
XX that are useful in diagnostic compositions for identifying T-cell
XX activating or co-stimulating compounds, or for identifying inhibitors of
XX T-cell activation and stimulation. TIRC7 peptides or polypeptides,
XX vectors and antibodies are useful in the preparation of pharmaceutical
XX compositions for treatment of acute and chronic diseases involving T-cell
XX activation and Th1 and Th2 immune response, for the treatment of acute
XX and chronic rejection of allo- and xeno organ transplants and bone marrow
XX transplantation, for the treatment of rheumatoid arthritis, lupus
XX erythematosus, multiple sclerosis, encephalitis, vasculitis, diabetes
XX mellitus, pancreatitis, gastritis, thyroiditis, for the treatment of
XX malign disorders of T, B or NK cells, for the treatment of asthma,
XX lepramatosi, Helicobacter pylori associated gastritis, or for the
XX treatment of skin, adrenal or lung tumors (claimed)

XX Sequence 614 AA;

Query Match 39.5%; Score 325; DB 2; Length 614;
Best Local Similarity 99.8%; Pred. No. 3.5e-309;
Matches 445; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 378 YGGRYQEVNPAPYTIITFPFLFAVMFGDVGHGLLMFLFALAMVLAENRPAVKAQNEIW 437
DB 162 YGGRYQEVNPAPYTIITFPFLFAVMFGDVGHGLLMFLFALAMVLAENRPAVKAQNEIW 221

QY 438 QTFPRGRYLLMLGLFSIYTGFIYNECFSRATSIFFPSGWSVAAMANGSGWSDAFLAQHTM 497
DB 222 QTFPRGRYLLMLGLFSIYTGFIYNECFSRATSIFFPSGWSVAAMANGSGWSDAFLAQHTM 281

QY 498 LTLDPNVTGVFLGYPFGIDPIWLSAANHLSFLNSFKMKMSVILGVVHMAFGVVLGVFNH 557
DB 282 LTLDPNVTGVFLGYPFGIDPIWLSAANHLSFLNSFKMKMSVILGVVHMAFGVVLGVFNH 341

QY 558 VFHQQRHRLLETLPELTFLGLFGYLVFLVIYKWLVCWAARAAS-PSILIHFINMFLFS 616
DB 342 VFHQQRHRLLETLPELTFLGLFGYLVFLVIYKWLVCWAARAASPSILIHFINMFLFS 401

QY 617 HSPSNRLLYPRQEVVQATLVVLAAMVPIILLGTPLHLHRRRRLRRPADROEENKAG 676
DB 402 HSPSNRLLYPRQEVVQATLVVLAAMVPIILLGTPLHLHRRRRLRRPADROEENKAG 461

QY 677 LLDLPDASVNGSSDEEKAGLDDEEAELVPSEVLHMQAIHTIEFCIGCVSNTASYLRL 736
DB 462 LLDLPDASVNGSSDEEKAGLDDEEAELVPSEVLHMQAIHTIEFCIGCVSNTASYLRL 521

QY 737 WALSLAHAQLSEVLWAMVMRIGLGLREVGVAAVVLVPIFAAFVMTVAILLVMEGLSAP 796
DB 522 WALSLAHAQLSEVLWAMVMRIGLGLREVGVAAVVLVPIFAAFVMTVAILLVMEGLSAP 581

QY 797 LHALLHWWBFQNKFSYGTGYKLSPP 822
DB 582 LHALLHWWBFQNKFSYGTGYKLSPP 607

RESULT 13

ABR43098
ID ABR43098 standard; protein; 614 AA.

XX
AC ABR43098;

XX
DT 03-JUL-2003 (first entry)

XX
DE Human T-cell immune response cDNA 7 (TIRC7) protein SEQ ID NO:5.

XX
KW Human; human leukocyte antigen; HLA; T-cell immune response cDNA 7;
KW TIRC7; immune response; immune disease; immunosuppressive; antiallergic;
KW antimicrobial; antibacterial; cytostatic; vulnerary.

XX
OS Homo sapiens.

XX
PN WO2003025000-A2.

XX
PD 27-MAR-2003.

XX
PF 17-SEP-2002; 2002WO-EF010440.

XX
PR 17-SEP-2001; 2001US-0322895P.

XX
PR 17-SEP-2001; 2001US-0322896P.

XX
PA (UTKU/) UTKU N.

XX
PI Utku N;

XX
DR WPI; 2003-381482/36.

XX
PT Peptide useful for the preparation of a pharmaceutical composition for
PT inhibition of an immune response comprises fragment of the amino acid
PT sequence.

XX
PS Claim 5; Fig 4A; 40pp; English.

XX

CC The present invention describes a peptide (P1) comprising a fragment of
CC the amino acid sequence from human leukocyte associated antigen (HLA)
CC class II alpha 2 chain (see ABR43094) or T cell immune response cDNA7
CC (TIRC7) protein (see ABR43098). Also described is a composition
CC comprising at least one (P1). (P1) has immunosuppressive, antiallergic,
CC antimicrobial, antibacterial, cytostatic and vulnerary activities, and
CC can be used as an inhibitor of the proliferation of peripheral blood
CC mononuclear cells (PBMCs), and as an immune response modulator. (P1) can
CC be used for the preparation of a pharmaceutical composition for the
CC inhibition of an immune response. (P1) can also be used for treating
CC graft versus host disease, autoimmune diseases, allergic diseases,
CC infectious diseases, sepsis and tumours. (P1) can be used for the
CC improvement of wound healing, and for inducing or maintaining immune
CC unresponsiveness. The present sequence represents the human TIRC7
CC protein, which is used in the exemplification of the present invention

XX
SQ Sequence 614 AA;

Query Match 39.5%; Score 325; DB 6; Length 614;
Best Local Similarity 99.8%; Pred. No. 3.5e-309;
Matches 445; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 378 YGGRYQEVNPAPYTIITFPFLFAVMFGDVGHGLLMFLFALAMVLAENRPAVKAQNEIW 437
DB 162 YGGRYQEVNPAPYTIITFPFLFAVMFGDVGHGLLMFLFALAMVLAENRPAVKAQNEIW 221

QY 438 QTFPRGRYLLMLGLFSIYTGFIYNECFSRATSIFFPSGWSVAAMANGSGWSDAFLAQHTM 497
DB 222 QTFPRGRYLLMLGLFSIYTGFIYNECFSRATSIFFPSGWSVAAMANGSGWSDAFLAQHTM 281

QY 498 LTLDPNVTGVFLGYPFGIDPIWLSAANHLSFLNSFKMKMSVILGVVHMAFGVVLGVFNH 557
DB 282 LTLDPNVTGVFLGYPFGIDPIWLSAANHLSFLNSFKMKMSVILGVVHMAFGVVLGVFNH 341

QY 558 VFHQQRHRLLETLPELTFLGLFGYLVFLVIYKWLVCWAARAAS-PSILIHFINMFLFS 616
DB 342 VFHQQRHRLLETLPELTFLGLFGYLVFLVIYKWLVCWAARAASPSILIHFINMFLFS 401

QY 617 HSPSNRLLYPRQEVVQATLVVLAAMVPIILLGTPLHLHRRRRLRRPADROEENKAG 676
DB 402 HSPSNRLLYPRQEVVQATLVVLAAMVPIILLGTPLHLHRRRRLRRPADROEENKAG 461

QY 677 LLDLPDASVNGSSDEEKAGLDDEEAELVPSEVLHMQAIHTIEFCIGCVSNTASYLRL 736
DB 462 LLDLPDASVNGSSDEEKAGLDDEEAELVPSEVLHMQAIHTIEFCIGCVSNTASYLRL 521

QY 737 WALSLAHAQLSEVLWAMVMRIGLGLREVGVAAVVLVPIFAAFVMTVAILLVMEGLSAP 796
DB 522 WALSLAHAQLSEVLWAMVMRIGLGLREVGVAAVVLVPIFAAFVMTVAILLVMEGLSAP 581

QY 797 LHALLHWWBFQNKFSYGTGYKLSPP 822
DB 582 LHALLHWWBFQNKFSYGTGYKLSPP 607

RESULT 14

ABP41524

ID ABP41524 standard; protein; 643 AA.

XX
AC ABP41524;

XX
DT 22-AUG-2002 (first entry)

XX
DE Human ovarian antigen HTAES83, SEQ ID NO:2656.

XX
KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;

XX
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;

KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antinflammatory; gynaecological; reproductive; chromosome 11q13.4-13.5.
 XX
 OS Homo sapiens.
 PN WO200200677-A1.
 XX
 XX
 PD 03-JAN-2002.
 XX
 XX
 PF 07-JUN-2001; 2001WO-US019569.
 XX
 XX
 PR 07-JUN-2000; 2000US-0209467P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX
 PI Birse CE, Rosen CA;
 XX N-PSDB; ABQ54601.
 DR
 DR
 XX
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
 PT cancer), immune disorders, cardiovascular disorders and neurological
 PT diseases.
 XX
 XX
 PS Claim 11; SEQ ID NO 2656; 2922pp; English.
 XX
 XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 643 AA;
 SQ

Query Match 35.4%; Score 291; DB 5; Length 643;
 Best Local Similarity 99.6%; Pred. No. 8.3e-276;
 Matches 491; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 110 DVRGNQALRAQLHQLHAAVLROGHEPQLAAHTDGAERTPLLPAGPGHQLRVNF 169
 DB 78 DVRGNQALRAQLHQLHAAVLROGHEPQLAAHTDGAERTPLLPAGPGHQLRVNF 137
 QY 170 VAGAVEPHKAPALERLLWRACRGFLIASPRELEQPLEHPVTGEPATMTFLISYGEQIG 229
 DB 138 VAGAVEPHKAPALERLLWRACRGFLIASPRELEQPLEHPVTGEPATMTFLISYGEQIG 197
 QY 230 QKIRKITDCFHCHVFFLQOEERLALQLOOQQSQELQEVLTETRFSLQVLRVQLL 289

Db 198 QKIRKITDCFHCHVFFLQOEERLALQLOOQQSQELQEVLTETRFSLQVLRVQLL 257
 QY 290 PFGQVQVHKMAVYLALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLDSSMEEGVSAVA 349
 Db 258 PFGQVQVHKMAVYLALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLDSSMEEGVSAVA 317
 QY 350 HRIPCRDMPPTLIRNRTFTASFGQIVDYGVCGRYQEVNPAPYTIITFPFLFAVMFGDVGH 409
 Db 318 XRIPCRDMPPTLIRNRTFTASFGQIVDYGVCGRYQEVNPAPYTIITFPFLFAVMFGDVGH 377
 QY 410 GLLMFLFALAMVLAENRPAVKAQNEIMOTFFRGYLLLLMGLFSIYTGFIYNECSRAT 469
 Db 378 GLLMFLFALAMVLAENRPAVKAQNEIMOTFFRGYLLLLMGLFSIYTGFIYNECSRAT 437
 QY 470 SIFPSGWSVAAMANQSGMSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWLSLAANHLSF 529
 Db 438 SIFPSGWSVAAMANQSGMSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWLSLAANHLSF 497
 QY 530 LNSFKMKMSVILGVVHMAFGVVLGVFNHVFQGRHLLLETLPELTFLGLGFLYVFLVI 589
 Db 498 LNSFKMKMSVILGVVHMAFGVVLGVFNHVFQGRHLLLETLPELTFLGLGFLYVFLVI 557
 QY 590 YKWLVCVMAARAAS 602
 Db 558 YKWLVCVMAARAAS 570

RESULT 15
 ADD27652
 ID ADD27652 standard; protein; 124 AA.
 XX
 AC ADD27652;
 XX
 DT 15-JAN-2004 (first entry)
 XX
 DE Human adipocyte Selected Interacting Domain (SID) prey protein #1109.
 XX
 KW Adipocyte; protein-protein interaction; protein complex;
 KW bait-prey complex; Selected Interacting Domain; SID; drug screening;
 KW drug discovery; metabolic disease; obesity; lipodystrophy;
 KW diabetes mellitus; type 2; non-insulin dependent; NIDDM;
 KW adipogenesis modulation; gene therapy; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200290544-A2.
 XX
 PD 14-NOV-2002.
 XX
 PF 03-MAY-2002; 2002WO-EP006333.
 XX
 PR 04-MAY-2001; 2001US-0288885P.
 XX
 XX (HYBR-) HYBRIGENICS.
 PA (LYNX-) LYNX THERAPEUTICS INC.
 XX
 XX Legrain P, Whiteside S, Mao J, Khrebukova I, Luo S;
 XX
 DR WPI; 2003-111975/10.
 DR N-PSDB; ADD27651.
 XX
 PT New complex of bait and prey between two polypeptides or polynucleotides
 PT encoding the two polypeptides of adipocytes, useful for selecting a
 PT modulating compound that inhibits or activates protein-protein
 PT interactions.
 XX
 PS Claim 6; SEQ ID NO 1109; 232pp; English.
 XX
 CC The invention relates to a bait-prey complex between two adipocyte
 CC polypeptides, or between two polynucleotides encoding adipocyte
 CC polypeptides. The invention also relates to selected interacting Domain
 CC (SID) polypeptides which interact with selected bait polypeptides;

```

CC polynucleotides encoding SID polypeptides; vectors comprising SID
CC polynucleotides; recombinant host cells comprising an adipocyte
CC polynucleotide or a SID-encoding vector; a method of selecting for a
CC compound which modulates interactions between adipocyte polypeptides;
CC adipocyte modulator compounds identified using the method; a
CC pharmaceutical composition comprising an adipocyte modulator, or a SID-
CC encoding vector or host cell; and a protein chip comprising adipocyte
CC bait polypeptides. The bait-prey complexes of the invention are useful
CC for selecting a compound that inhibits or activates protein-protein
CC interactions between adipocyte polypeptides. the modulatory compounds
CC identified can be used in the treatment of metabolic diseases such as
CC obesity, lipodystrophy and type 2 diabetes mellitus, and in the
CC modulation of adipogenesis. The present sequence represents a
CC specifically claimed adipocyte SID prey polypeptide of the invention.
XX
SQ Sequence 124 AA;

Query Match 15.1%; Score 124; DB 7; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.3e-112;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 606 LIHFINFLFESHSPNRLLYPRQEVQATLVLLALAMPVILLGCTPLHLLHRRRLRRR 665
Db 1 LIHFINFLFESHSPNRLLYPRQEVQATLVLLALAMPVILLGCTPLHLLHRRRLRRR 60

QY 666 PADROENKAGLLDLPDASVNGWSSDEKAGGLDDEEAELVPSEVLHQAIHTIECLG 725
Db 61 PADROENKAGLLDLPDASVNGWSSDEKAGGLDDEEAELVPSEVLHQAIHTIECLG 120

QY 726 CVSN 729
Db 121 CVSN 124

```

Search completed: June 29, 2006, 13:16:50
Job time : 197 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2006, 13:17:09 ; Search time 46 Seconds
(without alignments)
1719.352 Million cell updates/sec

Title: US-10-783-519-2
Perfect score: 822
Sequence: 1 MGSMPRSEVALVQLFLPTA.....HWVEFQNKFGYGTGKLSDF 822

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 1

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 80:*

1: Pirl:.*
2: Pirl:.*
3: Pirl:.*
4: Pirl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 27 | 3.3 | 163 | 2 T46449 | hypothetical prote |
| 2 | 27 | 3.3 | 838 | 2 A54163 | vacuolar ATPase (E |
| 3 | 24 | 2.9 | 838 | 2 B38656 | vacuolar proton pu |
| 4 | 22 | 2.7 | 873 | 2 T16282 | hypothetical prote |
| 5 | 22 | 2.7 | 1236 | 2 T19492 | hypothetical prote |
| 6 | 20 | 2.4 | 775 | 2 T46719 | probable vacuolar |
| 7 | 20 | 2.4 | 843 | 2 T06068 | probable proton pu |
| 8 | 20 | 2.4 | 855 | 2 JH0287 | immune regulatory |
| 9 | 19 | 2.3 | 935 | 2 S15795 | vacuolar proton pu |
| 10 | 18 | 2.2 | 840 | 1 A42970 | H+-exporting ATPas |
| 11 | 18 | 2.2 | 865 | 2 T18565 | probable H+-export |
| 12 | 18 | 2.2 | 890 | 2 S54554 | H+-exporting ATPas |
| 13 | 17 | 2.1 | 780 | 2 H84685 | probable vacuolar |
| 14 | 17 | 2.1 | 831 | 2 H84600 | probable vacuolar |
| 15 | 14 | 1.7 | 805 | 2 T37787 | probable vacuolar |
| 16 | 8 | 1.0 | 86 | 2 B82006 | hypothetical prote |
| 17 | 8 | 1.0 | 169 | 2 G71944 | hypothetical prote |
| 18 | 8 | 1.0 | 171 | 2 E64655 | hypothetical prote |
| 19 | 8 | 1.0 | 177 | 2 F87441 | hypothetical prote |
| 20 | 8 | 1.0 | 187 | 2 S23595 | embryonic fibrobla |
| 21 | 8 | 1.0 | 192 | 2 S54407 | embryonic fibrobla |
| 22 | 8 | 1.0 | 202 | 2 E87046 | probable lipoprote |
| 23 | 8 | 1.0 | 210 | 2 T41025 | hypothetical prote |
| 24 | 8 | 1.0 | 248 | 2 G83694 | hypothetical prote |
| 25 | 8 | 1.0 | 249 | 2 A70786 | probable Cobs - My |
| 26 | 8 | 1.0 | 255 | 2 G69533 | hypothetical prote |
| 27 | 8 | 1.0 | 263 | 2 T23473 | hypothetical prote |
| 28 | 8 | 1.0 | 324 | 2 T08729 | RING zinc finger p |
| 29 | 8 | 1.0 | 342 | 2 AF2177 | glucokinase [impor |

30 8 1.0 382 2 D84766 hypothetical prote
31 8 1.0 389 2 AD0271 conserved hypothet
32 8 1.0 442 2 D84224 heme synthase [imp
33 8 1.0 447 2 AF3313 osmolarity sensor
34 8 1.0 606 2 F71330 probable DNA repai
35 8 1.0 693 2 AF2357 hypothetical prote
36 8 1.0 758 2 A83121 probable two-compo
37 8 1.0 930 2 E96600 protein Fl4J16.20
38 7 0.9 36 1 JU0352 4.6K protein - Chl
39 7 0.9 93 2 I84714 H2-1A-beta-2 anti
40 7 0.9 93 2 I54421 MHC RT1.B-beta2 -
41 7 0.9 106 2 S66735 probable membrane
42 7 0.9 116 1 C64423 conserved hypothet
43 7 0.9 116 1 E69393 probable erp prot
44 7 0.9 118 2 H70378 hypothetical prote
45 7 0.9 134 2 A59055 phospholipase A2 (

ALIGNMENTS

RESULT 1

T46449
hypothetical protein DKFPz434H202.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46449
R;Ansoerge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23028
A;Accession: T46449
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-163 <AAA>
A;Cross-references: UNIPROT:Q93050; UNIPARC:UPI000016ACB4; EMBL:AL137683
A;Experimental source: adult testis; Clone DKFPz434H202
C;Genetics:
A;Note: DKFPz434H202.1
C;Superfamily: vacuolar ATP synthase 95K chain

Query Match 3.3%; Score 27; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 1e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 790 MEGLSAFLHALRLHWVEFQNKFGYSGTG 816
DB / 119 MEGLSAFLHALRLHWVEFQNKFGYSGTG 145

RESULT 2

A54163
vacuolar ATPase (EC 3.6.1.1-) 116k chain long form - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 09-Jul-2004
C;Accession: A54163
R;Peng, S.B.; Crider, B.P.; Xie, X.S.; Stone, D.K.
J. Biol. Chem. 269, 17262-17266, 1994
A;Title: Alternative mRNA splicing generates tissue-specific isoforms of 116-kDa polype
A;Reference number: A54163; MUID:94274722; PMID:8006034
A;Accession: A54163
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-838 <PEN>
A;Cross-references: UNIPROT:Q29466; UNIPARC:UPI0000138CEA; GB:L31770; NID:G469231; PDI
C;Superfamily: vacuolar ATP synthase 95K chain
C;Keywords: alternative splicing; hydrolase

Query Match 3.3%; Score 27; DB 2; Length 838;
Best Local Similarity 100.0%; Pred. No. 4.2e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 790 MEGLSAFLHALRLHWVEFQNKFGYSGTG 816
DB / 119 MEGLSAFLHALRLHWVEFQNKFGYSGTG 145

Db 794 MEGLSAFLHALRLHWVFQNKFGTGTG 820

RESULT 3
B38656
vacuolar proton pump 116K chain - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 21-Feb-1992 #sequence_revision 21-Feb-1992 #text_change 09-Jul-2004
C:Accession: B38656
R:Perin, M.S.; Fried, V.A.; Stone, D.K.; Xie, X.S.; Suedhof, T.C.
J. Biol. Chem. 266, 3877-3881, 1991
A:Title: Structure of the 116-kDa polypeptide of the clathrin-coated vesicle/synaptic vesicle
A:Reference number: A38656; MUID:91139686; PMID:1704894
A:Accession: B38656
A:Status: preliminary
A:Accession: T19492
A:Molecule type: mRNA
A:Residues: 1-838 <PER>
A:Cross-references: UNIPROT:P25286; UNIPARC:UPI0000138CD; GB:M58758; NID:G206429; PIDN:
C:Superfamily: vacuolar ATP synthase 95K chain
C:Keywords: alternative splicing; transmembrane protein

Query Match 2.9%; Score 24; DB 2; Length 838;
Best Local Similarity 100.0%; Pred. No. 4.3e-15; Mismatches 0; Indels 0; Gaps 0;
Matches 24; Conservative 0

QY 728 SNTASYLRWLWLSLAHAQLSEVLW 751
|||||
Db 734 SNTASYLRWLWLSLAHAQLSEVLW 757
|||||

RESULT 4
T16282
hypothetical protein F35H10.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16282
R:Wossner, J.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F35H10.
A:Reference number: Z18490
A:Accession: T16282
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-873 <WOS>
A:Cross-references: UNIPROT:Q20072; UNIPARC:UPI000008163D; EMBL:U40934; NID:G1072149; PI
C:Genetics:
A:Gene: CESP:F35H10.4
A:Introns: 39/3; 709/3; 785/1; 843/2
C:Superfamily: vacuolar ATP synthase 95K chain

Query Match 2.7%; Score 22; DB 2; Length 873;
Best Local Similarity 100.0%; Pred. No. 4.5e-13; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0

QY 730 TASYLRWLWLSLAHAQLSEVLW 751
|||||
Db 770 TASYLRWLWLSLAHAQLSEVLW 791
|||||

RESULT 5
T19492
hypothetical protein C26H9A.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T19492
R:Smyle, R.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z19131
A:Accession: T19492
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1236 <WIL>
A:Cross-references: UNIPROT:Q9XTS8; UNIPARC:UPI000017B7C6; EMBL:Z99169; PIDN:CAB16306.1;

A:Experimental source: clone C26H9A
C:Genetics:
A:Gene: CESP:C26H9A.1
A:Map position: 4
A:Introns: 50/2; 107/2; 119/3; 151/1; 209/2; 327/3; 432/3; 451/3; 482/2; 524/3; 552/2;
Query Match 2.7%; Score 22; DB 2; Length 1236;
Best Local Similarity 100.0%; Pred. No. 6.1e-13; Mismatches 0; Indels 0; Gaps 0;
Matches 22; Conservative 0

QY 385 EVNPAPYTIITPPFLFAVMFGD 406
|||||
Db 700 EVNPAPYTIITPPFLFAVMFGD 721
|||||

RESULT 6
T46719
probable vacuolar ATPase (EC 3.6.1.-) proton pump chain 116K [imported] - Leishmania m.
C:Species: Leishmania major
C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004
C:Accession: T46719
R:Voickart, G.; Ivens, A.C.; Lawson, D.; Quail, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, December 1999
A:Reference number: Z23137
A:Accession: T46719
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-775 <VOL>
A:Cross-references: UNIPROT:Q9UI50; UNIPARC:UPI000007C99D; EMBL:AL121861; PIDN:CAB58384
A:Experimental source: strain Friedlin
C:Genetics:
A:Note: L4326.08
C:Superfamily: vacuolar ATP synthase 95K chain
C:Keywords: hydrolase

Query Match 2.4%; Score 20; DB 2; Length 775;
Best Local Similarity 100.0%; Pred. No. 4.1e-11; Mismatches 20; Conservative 0; Indels 0; Gaps 0;
Matches 20

QY 724 LGCVSNTASYLRWLWLSLAH 743
|||||
Db 670 LGCVSNTASYLRWLWLSLAH 689
|||||

RESULT 7
T06068
probable proton pump F19H22.180 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C:Accession: T06068
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z15184
A:Accession: T06068
A:Molecule type: DNA
A:Residues: 1-843 <BEV>
A:Cross-references: UNIPROT:Q9SVI5; UNIPARC:UPI00000AAT64; EMBL:AL035679; GSPDB:GN00062
A:Experimental source: cultivar Columbia; BAC clone F19H22
C:Genetics:
A:Gene: ATP:F19H22.180
A:Map position: 4
A:Introns: 51/3; 66/3; 107/3; 143/3; 179/3; 237/3; 283/3; 319/3; 355/3; 405/2; 456/3; 5
C:Superfamily: vacuolar ATP synthase 95K chain

Query Match 2.4%; Score 20; DB 2; Length 843;
Best Local Similarity 100.0%; Pred. No. 4.4e-11; Mismatches 20; Conservative 0; Indels 0; Gaps 0;
Matches 20

QY 793 LSAPFLHALRLHWVFQNKFY 812
|||||
Db 804 LSAPFLHALRLHWVFQNKFY 823
|||||

RESULT 8

JH0287
immune regulatory protein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
A;Accession: JH0287
R;Lee, C.K.; Ghoshal, K.; Beaman, K.D.
Mol Immunol. 27, 1137-1144, 1990
A;Title: Cloning of a cDNA for a T cell produced molecule with a putative immune regulator
A;Reference number: JH0287; MUID:91061805; PMID:2247090
A;Accession: JH0287
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-855 <LEE>
A;Cross-references: UNIPROT:P15920; UNIPARC:UPI0000161FAB; GB:X55184; NID:G52754; PIDN:C
C;Superfamily: vacuolar ATP synthase 95K chain

Query Match 2.4%; Score 20; DB 2; Length 855;
Best Local Similarity 100.0%; Pred. No. 4.5e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 728 SNTASYLRRLWALSIAHAQLS 747
|||||
Db 748 SNTASYLRRLWALSIAHAQLS 767
|||||

RESULT 9

S15795
vacuolar proton pump homolog - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
A;Accession: S15795
R;Craxton, M.; Ainscough, R.; Coulson, A.; Dear, S.; Du, Z.; Durbin, R.; Green, P.; Hall
Staden, R.; Sulston, J.; Thierry-Mieg, J.; Thomas, K.; Waterston, R.; Wilson, R.
submitted to the EMBL Data Library, May 1991
A;Reference number: S15786
A;Accession: S15795
A;Molecule type: DNA
A;Residues: 1-935 <CRA>
A;Cross-references: UNIPROT:P30628; UNIPARC:UPI0000175EED; EMBL:Z11115; NID:G6953; PID:G
C;Genetics:
A;Introns: 51/3; 110/3; 153/3; 194/2; 264/2; 682/3; 723/3; 770/3; 844/1; 901/2
C;Superfamily: vacuolar ATP synthase 95K chain
C;Keywords: transmembrane protein

Query Match 2.3%; Score 19; DB 2; Length 935;
Best Local Similarity 100.0%; Pred. No. 4.9e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 790 MEGLSAPFLHALRLHWVEFQ 808
|||||
Db 887 MEGLSAPFLHALRLHWVEFQ 905
|||||

RESULT 10

A42970
H+-exporting ATPase (EC 3.6.3.6) 95K chain, vacuolar - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein O5430c; protein YOR270C
C;Species: Saccharomyces cerevisiae
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
A;Accession: A42970; S67167; S67172; S72041
R;Manolson, M.F.; Proteau, D.; Preston, R.A.; Stenbit, A.; Roberts, B.T.; Hoyt, M.A.; P
J. Biol. Chem. 267, 14294-14303, 1992
A;Title: The VPH1 gene encodes a 95-kDa integral membrane polypeptide required for in vi
A;Reference number: A42970; MUID:92332542; PMID:1385813
A;Accession: A42970
A;Molecule type: DNA
A;Residues: 1-840 <MAN>
A;Cross-references: UNIPROT:P32563; UNIPARC:UPI0000053075; GB:M89778; NID:G173172; PIDN:
A;Experimental source: strain X2180-1b; vacuolar acidification-defective mutants
R;Jauniaux, J.C.; Poirey, R.
submitted to the Protein Sequence Database, July 1996

A;Reference number: S67143
A;Accession: S67167
A;Molecule type: DNA
A;Residues: 1-840 <JAU>
A;Cross-references: UNIPARC:UPI0000053075; EMBL:Z75178; NID:G1420605; PIDN:CAA99494.1;
A;Experimental source: strain S288C
R;Cheret, G.; Sor, F.
submitted to the Protein Sequence Database, July 1996

A;Reference number: S67169
A;Accession: S67172
A;Molecule type: DNA
A;Residues: 1-840 <CHE>
A;Cross-references: UNIPARC:UPI0000053075; EMBL:Z75178; NID:G1420605; PIDN:CAA99494.1;
A;Experimental source: strain S288C
R;Cheret, G.; Bernardi, A.; Sor, F.
Yeast 12, 1059-1064, 1996

A;Title: DNA sequence analysis of the VPH1-SNF2 region on chromosome XV of Saccharomyces
A;Reference number: S72039; MUID:97051594; PMID:8896271
A;Accession: S72041
A;Molecule type: DNA
A;Status: nucleic acid sequence not shown; translation not shown
A;Residues: 1-840 <CHW>
A;Cross-references: UNIPARC:UPI0000053075; EMBL:X85633; NID:G1279694; PIDN:CAA61776.1;
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995

C;Genetics:
A;Gene: SGD:VPH1; MIPS:YOR270C
A;Cross-references: MIPS:YOR270c; SGD:S0005796
A;Map position: 15R
C;Function:
A;Description: hydrogen ion transport; hydrolase; required for assembly and activity of
C;Superfamily: vacuolar ATP synthase 95K chain
C;Keywords: ATP; glycoprotein; hydrogen ion transport; hydrolase; membrane-associated
F;407-441/Domain: transmembrane #status predicted <TM1>
F;457-478/Domain: transmembrane #status predicted <TM2>
F;539-558/Domain: transmembrane #status predicted <TM3>
F;565-591/Domain: transmembrane #status predicted <TM4>
F;635-656/Domain: transmembrane #status predicted <TM5>
F;733-795/Domain: transmembrane #status predicted <TM6>

Query Match 2.2%; Score 18; DB 1; Length 840;
Best Local Similarity 100.0%; Pred. No. 4.5e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 730 TASYLRRLWALSIAHAQLS 747
|||||
Db 730 TASYLRRLWALSIAHAQLS 747
|||||

RESULT 11

T18565
probable H+-exporting ATPase (EC 3.6.3.6) vacuolar [similarity] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
A;Accession: T18565; T19177
R;Mortimore, B.
submitted to the EMBL Data Library, December 1998

A;Reference number: Z18977
A;Accession: T18565
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-865 <WIL>
A;Cross-references: UNIPROT:Q17660; UNIPARC:UPI000007E97E; EMBL:AL031269; PIDN:CAA20334
A;Experimental source: clone VW02B12L
R;Coles, L.
submitted to the EMBL Data Library, August 1995

A;Reference number: Z19084
A;Accession: T19177
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-865 <WIL>
A;Cross-references: UNIPARC:UPI000007E97E; EMBL:Z50872; PIDN:CAA90758.1; GSPDB:GN000020;
A;Experimental source: clone C05D12
C;Genetics:

A:Gene: CESP:VW02B12L.1
 A:Map position: 2
 A:Introns: 66/1; 181/2; 576/2; 648/3; 775/1
 C:Superfamily: vacuolar ATP synthase 95K chain
 C:Keywords: hydrolase

Query Match 2.2%; Score 18; DB 2; Length 865;
 Best Local Similarity 100.0%; Pred. No. 4.7e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 730 TASYLRLWALSLSHAQLS 747
 DB 760 TASYLRLWALSLSHAQLS 777
 |||||

RESULT 12

S54554 H+-exporting ATPase (EC 3.6.3.6) chain STV1 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YM9796.07; protein YMR054W

C:Species: Saccharomyces cerevisiae
 C:Date: 08-Jul-1995 #sequence_revision 11-Aug-1995 #text_change 09-Jul-2004
 C:Accession: S54554; A54081
 R:Devlin, K.; Churcher, C.M.
 submitted to the EMBL Data Library, May 1995

A:Reference number: S54548
 A:Accession: S54554

A:Molecule type: DNA

A:Residues: 1-890 <DEV>

A:Cross-references: UNIPROT:P37296; UNIPARC:UPI000013615B; EMBL:Z49703; NID:g817880; PID:R:Manolson, M.F.; Wu, B.; Proteau, D.; Tallon, L.A.; Roberts, B.T.; Hoyt, M.A.; Jones, J. Biol. Chem. 269, 14064-14074, 1994

A:Title: STV1 gene encodes functional homologue of 95-kDa yeast vacuolar H(+)-ATPase subunit
 A:Reference number: A54081; MUID:94245725; PMID:7514599

A:Accession: A54081

A:Molecule type: DNA

A:Residues: 1-804, 'E', 806-890 <MAN>

A:Cross-references: UNIPARC:UPI00001689CE; GB:U06465; NID:g460159; PIDN:AAA20596.1; PID:J. Biol. Chem. 269, 14064-14074, 1994

C:Genetics:

A:Gene: SGD:STV1

A:Cross-references: SGD:S0004658; MIPS:YMR054W

A:Map position: 13R

C:Function:

A:Description: hydrogen ion transport; hydrolase
 C:Superfamily: vacuolar ATP synthase 95K chain
 C:Keywords: ATP; hydrogen ion transport; hydrolase; transmembrane protein; yeast vacuole
 F:451-487/Domain: transmembrane #status predicted <TM1>
 F:509-525/Domain: transmembrane #status predicted <TM2>
 F:585-600/Domain: transmembrane #status predicted <TM3>
 F:614-637/Domain: transmembrane #status predicted <TM4>
 F:681-696/Domain: transmembrane #status predicted <TM5>
 F:833-851/Domain: transmembrane #status predicted <TM6>

Query Match 2.2%; Score 18; DB 2; Length 890;
 Best Local Similarity 100.0%; Pred. No. 4.7e-09;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 730 TASYLRLWALSLSHAQLS 747
 DB 790 TASYLRLWALSLSHAQLS 807
 |||||

RESULT 13

H84685 probable vacuolar proton-ATPase subunit [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C:Accession: H84685
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: H84685
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-780 <STO>
 A:Cross-references: UNIPROT:Q9SK06; UNIPARC:UPI00000A9C80; GB:AE002093; NID:g6598613;
 C:Genetics:
 A:Gene: At2g28520
 A:Map position: 2
 C:Superfamily: vacuolar ATP synthase 95K chain

Query Match 2.1%; Score 17; DB 2; Length 780;
 Best Local Similarity 100.0%; Pred. No. 4.2e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 727 VSNTASYLRLWALSLSAH 743
 DB 685 VSNTASYLRLWALSLSAH 701
 |||||

RESULT 14

H84600 probable vacuolar proton-ATPase subunit [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C:Accession: H84600

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: H84600

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-821 <STO>

A:Cross-references: UNIPROT:Q9SJT7; UNIPARC:UPI00000A1530; GB:AE002093; NID:g4567273;
 C:Genetics:

A:Gene: At2g21410

A:Map position: 2

C:Superfamily: vacuolar ATP synthase 95K chain

Query Match 2.1%; Score 17; DB 2; Length 821;
 Best Local Similarity 100.0%; Pred. No. 4.4e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 727 VSNTASYLRLWALSLSAH 743
 DB 721 VSNTASYLRLWALSLSAH 737
 |||||

RESULT 15

T37787 probable vacuolar atpase subunit - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T37787

R:Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 submitted to the EMBL Data Library, September 1995

A:Reference number: Z21746

A:Accession: T37787

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-805 <OLI>

A:Cross-references: UNIPROT:O13742; UNIPARC:UPI0000138CD1; EMBL:Z98529; PIDN:CAB11035
 A:Experimental source: strain 972h-; cosmid c16E8

C:Genetics:

A:Gene: SPDB:SPAC16E8.07c

A:Map position: 1

A:Introns: 14/3

C:Superfamily: vacuolar ATP synthase 95K chain

Query Match 1.7%; Score 14; DB 2; Length 805;
 Best Local Similarity 100.0%; Pred. No. 4.4e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 730 TASYRLWLALSLAH 743

Db 708 TASYRLWLALSLAH 721

Search completed: June 29, 2006, 13:22:56
Job time : 48 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2006, 13:13:44 ; Search time 308 Seconds
(without alignments)
2468.710 Million cell updates/sec

Title: US-10-783-519-2

Perfect score: 822

Sequence: 1 MGSMPRSEVALVQLFLPTA.....HWVEFQNKFSYGTGYKLSPF 822

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2849598 seqs, 925015592 residues

Word size : 1

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Uniprot 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------|
| 1 | 601 | 73.1 | 830 | 1 | VPP3_HUMAN |
| 2 | 545 | 66.3 | 830 | 2 | Q8WVC5_HUMAN |
| 3 | 153 | 18.6 | 218 | 2 | Q64F87_HUMAN |
| 4 | 60 | 7.3 | 60 | 2 | Q6QBN6_HUMAN |
| 5 | 58 | 7.1 | 125 | 2 | Q6QBN8_HUMAN |
| 6 | 46 | 5.6 | 61 | 2 | Q8TCH1_HUMAN |
| 7 | 40 | 4.9 | 430 | 2 | Q9CTA9_MOUSE |
| 8 | 40 | 4.9 | 834 | 2 | Q9JHF5_MOUSE |
| 9 | 40 | 4.9 | 834 | 2 | Q9JL12_MOUSE |
| 10 | 40 | 4.9 | 834 | 2 | Q9LW06_MOUSE |
| 11 | 40 | 4.9 | 834 | 2 | Q216B0_RAT |
| 12 | 38 | 4.6 | 420 | 2 | Q6P735_RAT |
| 13 | 35 | 4.3 | 45 | 2 | Q6QBN7_HUMAN |
| 14 | 34 | 4.1 | 822 | 2 | Q7ZVM7_BRAKE |
| 15 | 32 | 3.9 | 846 | 2 | Q6PA83_XENLA |
| 16 | 30 | 3.6 | 833 | 1 | VPP4_MOUSE |
| 17 | 30 | 3.6 | 837 | 2 | Q918C8_CHICK |
| 18 | 29 | 3.5 | 838 | 2 | Q4RSZ3_TETNG |
| 19 | 29 | 3.5 | 839 | 2 | Q7TIN9_TORMA |
| 20 | 29 | 3.5 | 840 | 2 | Q7TIN8_TORMA |
| 21 | 29 | 3.5 | 897 | 2 | Q4SJ89_TETNG |
| 22 | 27 | 3.3 | 831 | 2 | Q53ET5_HUMAN |
| 23 | 27 | 3.3 | 831 | 2 | Q53X12_HUMAN |
| 24 | 27 | 3.3 | 834 | 2 | Q6NY92_BRAKE |
| 25 | 27 | 3.3 | 837 | 1 | VPP1_HUMAN |
| 26 | 27 | 3.3 | 837 | 1 | VPP1_PONPY |
| 27 | 27 | 3.3 | 837 | 2 | Q5R5X1_PONPY |
| 28 | 27 | 3.3 | 837 | 2 | Q5R6N4_PONPY |
| 29 | 27 | 3.3 | 838 | 1 | VPP1_BOVIN |
| 30 | 27 | 3.3 | 838 | 2 | Q5CZHE_HUMAN |
| 31 | 26 | 3.2 | 95 | 2 | Q6QBN5_HUMAN |

| | | | | | |
|----|----|-----|-----|---|--------------|
| 32 | 26 | 3.2 | 852 | 2 | Q4DY50_TRYCR |
| 33 | 26 | 3.2 | 854 | 1 | VPP2_BOVIN |
| 34 | 25 | 3.0 | 169 | 2 | Q9JL11_MOUSE |
| 35 | 25 | 3.0 | 823 | 2 | Q5CZ26_XENTR |
| 36 | 25 | 3.0 | 827 | 2 | Q4RZB2_TETNG |
| 37 | 25 | 3.0 | 835 | 2 | Q4S964_TETNG |
| 38 | 24 | 2.9 | 113 | 2 | Q4RBL1_TETNG |
| 39 | 24 | 2.9 | 161 | 2 | Q5DI34_SCHJA |
| 40 | 24 | 2.9 | 236 | 2 | Q5C2K8_SCHJA |
| 41 | 24 | 2.9 | 483 | 2 | Q6AHY6_HUMAN |
| 42 | 24 | 2.9 | 513 | 2 | Q96N91_HUMAN |
| 43 | 24 | 2.9 | 779 | 2 | Q3TV98_MOUSE |
| 44 | 24 | 2.9 | 831 | 2 | Q8AVM5_XENLA |
| 45 | 24 | 2.9 | 832 | 2 | Q3TXT5_MOUSE |

ALIGNMENTS

RESULT 1
VPP3_HUMAN
ID VPP3_HUMAN STANDARD; PRS; 830 AA.
AC Q13458; O75877;
DT 01-DEC-2000, integrated into UniProtKB/Swiss-Prot.
DT 01-DEC-2000, sequence version 2.
DT 07-FEB-2006, entry version 49.
DE Vacuolar proton translocating ATPase 116 kDa subunit a isoform 3 (V-ATPase 116-kDa isoform a3) (Osteoclastic proton pump 116 kDa subunit) (OC-116 kDa) (OC116) (T-cell immune regulator 1) (T cell immune response cDNA7 protein) (TIRC7)
GN Names=TCIRG1; Synonyms=ATP6N1C, ATP6V0A3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM LONG).
RC TISSUE=Osteoclastoma;
RX MEDLINE=96158968; PubMed=8579597; DOI=10.1006/bbrc.1996.0145;
RA Li Y.-P., Chen W., Stashenko P.;
RT "Molecular cloning and characterization of a putative novel human osteoclast-specific 116-kDa vacuolar proton pump subunit.";
RL Biochem. Biophys. Res. Commun. 218:813-821(1996).
[2]
RP NUCLEOTIDE SEQUENCE (ISOFORM SHORT).
RA Utaku N., Heinemann T., Bulwin C.-G., Beinke S., Beato F., Randall J., Busconi L., Delphire E., Robertson E.R., Kojima R., Volk H.D., Milford E.L., Gullans S.R.;
RA Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
[3]
RP NUCLEOTIDE SEQUENCE (ISOFORM SHORT).
RX MEDLINE=99263502; PubMed=10329006; DOI=10.1006/geno.1999.5751;
RA Heinemann T., Bulwin G.C., Randall J., Schnieders B., Sandhoff K., Volk H.D., Milford E., Gullans S.R., Utaku N.;
RT "Genomic organization of the gene coding for TIRC7, a novel membrane protein essential for T cell activation.";
RL Genomics 57:398-406(1999).
[4]
RP VARIANTS OPTB1 ARG-405 AND LEU-444.
RX PubMed=11532986; DOI=10.1093/hmg/10.17.1767;
RA Sobacchi C., Frattini A., Orchard P., Porras O., Texcan I., Andolina M., Babul-Hirji R., Baric I., Canham N., Chitayat D., Dupuis-Girod S., Ellis I., Etzioni A., Faeth A., Fisher A., Gerritsen B., Gulino V., Horwitz E., Klamroth V., Lanino E., Mirolto M., Musio A., Matichis G., Nonomaya S., Notarangelo L.D., Ochs H.D., Superti Furga A., Valiaho J., van Hove J.L.K., Vihinen M., Vujic D., Vezzoni P., Villa A.;
RT "The mutational spectrum of human malignant autosomal recessive osteopetrosis.";
RL Hum. Mol. Genet. 10:1767-1773(2001).
[5]
RP VARIANT OPTB1 ARG-405.

Q4dy50 trypanosoma
Q37681 bos taurus
Q9JL11 mus musculus
Q5CZ26 xenopus tro
Q4RZB2 tetraodon n
Q4S964 tetraodon n
Q4RBL1 tetraodon n
Q5DI34 schistosoma
Q5C2K8 schistosoma
Q6AHY6 homo sapien
Q96N91 homo sapien
Q3TV98 mus musculus
Q8AVM5 xenopus lae
Q3TXT5 mus musculus

XX PubMed=12552563; DOI=10.1002/humu.10165;
RA Scineca J.-C., Quincey D., Perrinello H., Romatet D., Grosgeorge J.,
RA Gaudray P., Philip N., Fischer A., Carle G.F.;
RT "Novel mutations in the TCIRG1 gene encoding the a3 subunit of the
RT vacuolar proton pump in patients affected by infantile malignant
RT osteopetrosis";
RL Hum. Mutat. 21:151-157(2003).
RN [6].
RP VARIANTS OPTB1 PRO-141; ARG-405; ASN-462 DEL; ASN-517 AND ARG-775.
RX PubMed=15300850; DOI=10.1002/humu.20076;
RA Susani L., Pangrazio A., Sobacchi C., Taranta A., Mortier G.,
RA Savarirayan R., Villa A., Orchard P., Vezzoni P., Albertini A.,
RA Frattini A., Pagani F.;
RT "TCIRG1-dependent recessive osteopetrosis: mutation analysis,
RT functional identification of the splicing defects, and in vitro rescue
RT by U1 snRNA";
RL Hum. Mutat. 24:225-235(2004).
CC -I- FUNCTION: Part of the proton channel of V-ATPases (By similarity).
CC Seems to be directly involved in T cell activation.
CC -I- SUBUNIT: The V-ATPase is an heteromultimeric enzyme composed of at
CC least thirteen different subunits. It has a membrane peripheral V1
CC sector for ATP hydrolysis and an integral V0 for proton
CC translocation. The V1 sector contains subunits A-H, whereas V0
CC includes subunits a, d, c, c', and c''.
CC -I- SUBCELLULAR LOCATION: Membrane; multi-pass membrane protein (By
CC similarity).
CC -I- ALTERNATIVE PRODUCTS:
CC Event=alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=Q13488-1; Sequences=Displayed;
CC Name=Short;
CC IsoId=Q13488-2; Sequences=VSP 000345;
CC Note=No experimental confirmation available;
CC -I- TISSUE SPECIFICITY: The long isoform is highly expressed in
CC osteoclasts. The short isoform is highly expressed in thymus.
CC -I- DISEASE: Defects in TCIRG1 are a cause of autosomal recessive
CC osteopetrosis (OPTB1) [MIM:259700]; also called autosomal
CC recessive Albers-Schönberg disease. Osteopetrosis is a rare
CC genetic disease characterized by abnormally dense bone, due to
CC defective resorption of immature bone. The defect is an
CC heterogeneous disorder of bone metabolism, which, if untreated, has
CC a fatal outcome. It occurs in two forms: a severe autosomal
CC recessive form, occurring in utero, infancy, or childhood, and a
CC benign autosomal dominant form, occurring in adolescence or
CC adulthood. However, mild and atypical forms have also been
CC reported. It is not known whether these latter cases represent
CC separate nosological entities or a mild form of the recessive
CC conditions. However, given the large number of genes and gene
CC mutations associated with this condition a large spectrum of
CC clinical manifestation is expected as seen in animal studies. The
CC features of OPTB1 are macrocephaly, progressive deafness and
CC blindness, hepatosplenomegaly, and severe anemia beginning in
CC early infancy or in fetal life. Deafness and blindness are
CC generally thought to represent effects of pressure on nerves.
CC -I- SIMILARITY: Belongs to the V-ATPase 116 kDa subunit family.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC EMBL; U45285; AAA97878.1; -; mRNA.
CC EMBL; AF025374; AAC35742.1; -; mRNA.
CC EMBL; AF033033; AAD31081.2; -; Genomic DNA.
CC Ensembl; ENSG00000110719; Homo sapiens.
CC H-InvDB; HIX0009874; -;
CC HGNC; HGNC:11847; TCIRG1.
CC MIM; 259700; phenotype.
CC MIM; 604592; Gene.
CC DR GO; GO:0005887; C: integral to plasma membrane; TAS.
CC DR GO; GO:0005886; C: plasma membrane; TAS.
CC DR GO; GO:0005215; P: transporter activity; TAS.
CC DR GO; GO:0006968; P: cellular defense response; TAS.
CC DR GO; GO:0008284; P: positive regulation of cell proliferation; TAS.
CC DR GO; GO:0015992; P: proton transport; TAS.

DR InterPro: IPR002490; V_ATPase_sub116.
DR PANTHER: PTHR11629; V_ATPase_sub116; 1.
DR Pfam; Pfold96; V_ATPase_1; 1.
KW Alternative splicing; Disease mutation; Glycoprotein;
KW Hydrogen ion transport; Ion transport; Membrane; Transmembrane;
KW Transport.
FT CHAIN 1 830 Vacuolar proton translocating ATPase 116
FT KDa subunit a isoform 3.
FT /FTID=PRO_0000119218.
FT TOPO_DOM 1 397 Extracellular (Potential).
FT TRANSMEM 398 418 Potential.
FT TRANSMEM 445 465 Potential.
FT TRANSMEM 504 524 Potential.
FT TRANSMEM 538 558 Potential.
FT TRANSMEM 576 596 Potential.
FT TRANSMEM 636 656 Potential.
FT TRANSMEM 771 791 Potential.
FT CARBOHYD 41 41 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 483 483 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 503 503 N-linked (GlcNAc...) (Potential).
FT VARSPLIC 1 216 Missing (in isoform Short).
FT /FTID=VSP_000345.
FT VARIANT 141 141 A -> P (in OPTB1).
FT VARIANT 405 405 /FTID=VAR_020988.
FT VARIANT 444 444 G -> R (in OPTB1).
FT VARIANT 462 462 /FTID=VAR_019569.
FT VARIANT 517 517 R -> L (in OPTB1).
FT VARIANT 775 775 /FTID=VAR_019570.
FT CONFLICT 377 377 Missing (in OPTB1).
FT CONFLICT 603 603 /FTID=VAR_020989.
FT SEQUENCE 830 AA; 92998 MW; 62EAC9A22DC698B CRC64;
Query Match 73.1%; Score 601; DB 1; Length 830;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 821; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 MGSMFSEVALVQLFPTAAAYTCVRLGELGVFEFRDLNASVAFQRRFVVDVWRCE 60
Db 1 MGSMFSEVALVQLFPTAAAYTCVRLGELGVFEFRDLNASVAFQRRFVVDVWRCE 60
QY 61 LEKTFTLQEEVRRAGLVLPKPKRGLPAPPDRDLRIQETTERLAQELRDVGNQALRA 120
Db 61 LEKTFTLQEEVRRAGLVLPKPKRGLPAPPDRDLRIQETTERLAQELRDVGNQALRA 120
QY 121 QLHQQLHAANVLROGHEPQLAAAHDTGASERTPLLOAGPGPHQDLRVNPFVAGAVEPHKAP 180
Db 121 QLHQQLHAANVLROGHEPQLAAAHDTGASERTPLLOAGPGPHQDLRVNPFVAGAVEPHKAP 180
QY 181 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEOIGQKIRK1TDCPH 240
Db 181 ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEOIGQKIRK1TDCPH 240
QY 241 CHVFPFLQEEARLQALQQLQSQSQELQELGETERFLISQVLGRVLQLLPQGVQVHKMK 300
Db 241 CHVFPFLQEEARLQALQQLQSQSQELQELGETERFLISQVLGRVLQLLPQGVQVHKMK 300
QY 301 AVYLALNOCVSSTTHKCLIAEAWCSVRDLPALQELALROSSMBEGVSVAHRI1PCRDMPPT 360
Db 301 AVYLALNOCVSSTTHKCLIAEAWCSVRDLPALQELALROSSMBEGVSVAHRI1PCRDMPPT 360
QY 361 LIRTNRTASFGIVDRYGVGRYQVBNPAPYTIITFPFLFAVMFGDVGHLMLFALAM 420
Db 361 LIRTNRTASFGIVDRYGVGRYQVBNPAPYTIITFPFLFAVMFGDVGHLMLFALAM 420
QY 421 VLAENRPVAKAQNINWOTFFRGRYLLLMGLFISYTGFIYNECFSRATSIFFPSGMSVAA 480
Db 421 VLAENRPVAKAQNINWOTFFRGRYLLLMGLFISYTGFIYNECFSRATSIFFPSGMSVAA 480

QY 481 MANQSGWSDAFIAQHTMLTLDPNVTGVFLGYPGIDPIWLSLAANHLSFLNSFKMKMSVI 540
Db 481 MANQSGWSDAFIAQHTMLTLDPNVTGVFLGYPGIDPIWLSLAANHLSFLNSFKMKMSVI 540
QY 541 LGVVMAGVVLGVFNHVFQGRHLLLETLPELTLLGLFGYLVFLVIVYKWLVCWAARA 600
Db 541 LGVVMAGVVLGVFNHVFQGRHLLLETLPELTLLGLFGYLVFLVIVYKWLVCWAARA 600
QY 601 AS-PSILIHFINMPLFSGHSPSNRLLYPRQEVQVQATLVVLALAMVPILLGLTPLHLHRH 659
Db 601 ASAPSILIHFINMPLFSGHSPSNRLLYPRQEVQVQATLVVLALAMVPILLGLTPLHLHRH 660
QY 660 RLRRRRPADROENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLHQAIHT 719
Db 661 RLRRRRPADROENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLHQAIHT 720
QY 720 IEFCLGCVSNNTASYLRRLWALSALHAQSEVLWAMVMRIGLGLGREVGVAALVPIPAAF 779
Db 721 IEFCLGCVSNNTASYLRRLWALSALHAQSEVLWAMVMRIGLGLGREVGVAALVPIPAAF 780
QY 780 AVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPF 822
Db 781 AVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPF 823

RESULT 2

Q8WVC5 HUMAN
ID Q8WVC5 HUMAN PRELIMINARY; PRT; 830 AA.
AC Q8WVC5;
DT 01-MAR-2002, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2002, sequence version 1.
DT 21-FEB-2006, entry version 20.
DE T-cell, immune regulator 1, isoform a.
GN Names=TCIRG1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=lung and pancreas;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalek U., Smailus D.E.,
RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=pancreas;
RG NIH MGC Project;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=lung;
RG NIH MGC Project;

RESULT 3

RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC018133; AA18133.1; -; mRNA.
DR EMBL; BC032465; AA32465.1; -; mRNA.
DR Ensembl; ENSG00000110719; Homo sapiens.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR002490; V ATPase sub116.
DR PANTHER; PTHR11629; V ATPase sub116; 1.
DR Pfam; PF01456; V ATPase_1; 1.
SQ SEQUENCE 830.AA; 92968 MW; 50751B41B171D9D2 CRC64;

Query Match 66.3%; Score 545; DB 2; Length 830;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 765; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 57 RCELEKTFFTFLOEEVRRAGLVLPKGRPLPAPPRDLLRIQEBETERLAQELRDVRNQ 116
Db 57 RCELEKTFFTFLOEEVRRAGLVLPKGRPLPAPPRDLLRIQEBETERLAQELRDVRNQ 116
QY 117 ALRAQLHQLHAALVLRQGHPEQLAAHTDGAERTPLQAPGGPHODLRVNFVAGAVEP 176
Db 117 ALRAQLHQLHAALVLRQGHPEQLAAHTDGAERTPLQAPGGPHODLRVNFVAGAVEP 176
QY 177 HKPALERLILWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGQIGOKIRKIT 236
Db 177 HKPALERLILWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGQIGOKIRKIT 236
QY 237 DCFHCHVFPLQOEAEARLQALQOQSQBELQVGLGETERFLSQVLRVLLQPLPPGVQV 296
Db 237 DCFHCHVFPLQOEAEARLQALQOQSQBELQVGLGETERFLSQVLRVLLQPLPPGVQV 296
QY 297 HKMKAVYLALNQCSTVTHKCLIAEAWCSVRDIPALQEARLDSMEEGVSAVAHRIICRD 356
Db 297 HKMKAVYLALNQCSTVTHKCLIAEAWCSVRDIPALQEARLDSMEEGVSAVAHRIICRD 356
QY 357 MPPTLIRNRTFASFOGIVDVRVGVGRYQEVNRPAYTIITPPFLPFAVMFGVGHGLMLFLF 416
Db 357 MPPTLIRNRTFASFOGIVDVRVGVGRYQEVNRPAYTIITPPFLPFAVMFGVGHGLMLFLF 416
QY 417 ALAMVLAENRPVAKAAQNEIWQTFRRGRVYLLLMGLFSIYTGFIYNECFSRATSIIPSGW 476
Db 417 ALAMVLAENRPVAKAAQNEIWQTFRRGRVYLLLMGLFSIYTGFIYNECFSRATSIIPSGW 476
QY 477 SVAAANQSGWSDAFIAQHTMLTLDPNVTGVFLGYPGIDPIWLSLAANHLSFLNSFKMK 536
Db 477 SVAAANQSGWSDAFIAQHTMLTLDPNVTGVFLGYPGIDPIWLSLAANHLSFLNSFKMK 536
QY 537 MSVILGVVMAGVVLGVFNHVFQGRHLLLETLPELTLLGLFGYLVFLVIVYKWLVCW 596
Db 537 MSVILGVVMAGVVLGVFNHVFQGRHLLLETLPELTLLGLFGYLVFLVIVYKWLVCW 596
QY 597 ARAAAS-PSILIHFINMPLFSGHSPSNRLLYPRQEVQVQATLVVLALAMVPILLGLTPLHL 655
Db 597 ARAAASPSILIHFINMPLFSGHSPSNRLLYPRQEVQVQATLVVLALAMVPILLGLTPLHL 656
QY 656 HRHRRRLRRPADROENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLHQ 715
Db 657 HRHRRRLRRPADROENKAGLLDLPDASVNGWSSDEEKAGGLDDEEAEELVPSEVLHQ 716
QY 716 AHTHTIEFCLGCVSNNTASYLRRLWALSALHAQSEVLWAMVMRIGLGLGREVGVAALVPI 775
Db 717 AHTHTIEFCLGCVSNNTASYLRRLWALSALHAQSEVLWAMVMRIGLGLGREVGVAALVPI 776
QY 776 FFAFAVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPF 822
Db 777 FFAFAVMTVAILLVMEGLSAFLHALRLHWVEFQNKFSYGTGYKLSPF 823

Q64F87 HUMAN
 ID Q64F87_HUMAN PRELIMINARY; PRT; 218 AA.
 AC Q64F87
 DT 25-OCT-2004, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, sequence version 1.
 DE T-cell immune regulator 1 transcript variant 4 (Fragment).
 GN Name=TCIRG1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC Tissue=Galbladder;
 RX PubMed=15809087; DOI=10.1016/j.bbrc.2005.03.065;
 RA Smirnova A.S., Morgun A., Shulzhenko N., Silva I.D.,
 RA Gerbase-Delima M.;
 RT "Identification of new alternative splice events in the TCIRG1 gene in
 RT different human tissues";
 RL Biochem. Biophys. Res. Commun. 330:943-949(2005).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 DR EMBL; AY708388; AAU20795.1; -; mRNA.
 DR Ensembl; ENSG00000110719; Homo sapiens.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0015992; P:proton transport; IEA.
 DR InterPro; IPR002490; V_ATPase_sub116.
 DR PANTHER; PTHR11629; V_ATPase_sub116; 1.
 DR Pfam; PF01496; V_ATPase_I; 1.
 DR NON_TER 218 218
 SQ SEQUENCE 218 AA; 24383 MW; 55F80AB2B2A79DBB CRC64;
 Query Match 18.6%; Score 153; DB 2; Length 218;
 Best Local Similarity 100.0%; Pred. No. 1.7e-147;
 Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 57 RCELEKFTTFLQEVRRAGLVLPKGLPAPPDRDLRIQETTERLAQELRDVRGNQ 116
 DB 57 RCELEKFTTFLQEVRRAGLVLPKGLPAPPDRDLRIQETTERLAQELRDVRGNQ 116
 QY 117 ALRAQLHQLHAAVLROGHEPQLAAHTDGAERTLLQAPGGPHQDLRVNFVAGAVEP 176
 DB 117 ALRAQLHQLHAAVLROGHEPQLAAHTDGAERTLLQAPGGPHQDLRVNFVAGAVEP 176
 QY 177 HKAPALERLLMRACRGLFIASPRELEQPLEHPV 209
 DB 177 HKAPALERLLMRACRGLFIASPRELEQPLEHPV 209
 RESULT 4
 Q6QBN6 HUMAN
 ID Q6QBN6_HUMAN PRELIMINARY; PRT; 60 AA.
 AC Q6QBN6
 DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
 DT 05-JUL-2004, sequence version 1.
 DT 07-FEB-2006, entry version 8.
 DE T-cell immune regulator 1 transcript variant 6 (Fragment).
 GN Name=TCIRG1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC Tissue=Tonsil;
 RX PubMed=15809087; DOI=10.1016/j.bbrc.2005.03.065;
 RA Smirnova A.S., Morgun A., Shulzhenko N., Silva I.D.,
 RA Gerbase-Delima M.;
 RT "Identification of new alternative splice events in the TCIRG1 gene in
 RT different human tissues";
 RL Biochem. Biophys. Res. Commun. 330:943-949(2005).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 DR EMBL; AY548967; AAS59834.1; -; mRNA.
 DR Ensembl; ENSG00000110719; Homo sapiens.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0015992; P:proton transport; IEA.
 DR InterPro; IPR002490; V_ATPase_sub116.
 DR PANTHER; PTHR11629; V_ATPase_sub116; 1.
 DR Pfam; PF01496; V_ATPase_I; 1.
 DR NON_TER 125 125
 SQ SEQUENCE 125 AA; 14272 MW; C7AA604E66F54444 CRC64;
 Query Match 7.1%; Score 58; DB 2; Length 125;
 Best Local Similarity 100.0%; Pred. No. 3.7e-50;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 57 RCELEKFTTFLQEVRRAGLVLPKGLPAPPDRDLRIQETTERLAQELRDVRGN 114
 DB 57 RCELEKFTTFLQEVRRAGLVLPKGLPAPPDRDLRIQETTERLAQELRDVRGN 114
 RESULT 6

RT "Identification of new alternative splice events in the TCIRG1 gene in
 RT different human tissues";
 RL Biochem. Biophys. Res. Commun. 330:943-949(2005).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 DR EMBL; AY548969; AAS59836.1; -; mRNA.
 DR Ensembl; ENSG00000110719; Homo sapiens.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0015992; P:proton transport; IEA.
 DR InterPro; IPR002490; V_ATPase_sub116.
 DR PANTHER; PTHR11629; V_ATPase_sub116; 1.
 DR Pfam; PF01496; V_ATPase_I; 1.
 DR NON_TER 1 1
 SQ SEQUENCE 60 AA; 6545 MW; 5009B0C221BD4E96 CRC64;
 Query Match 7.3%; Score 60; DB 2; Length 60;
 Best Local Similarity 100.0%; Pred. No. 1.8e-52;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 459 FYINECFSRATSIIPSGWSVAAMANQSGMSDAFLAQHTMLTLDPNVTGVLGPPFGIDP 518
 DB 1 FYINECFSRATSIIPSGWSVAAMANQSGMSDAFLAQHTMLTLDPNVTGVLGPPFGIDP 518
 RESULT 5
 Q6QBN8 HUMAN
 ID Q6QBN8_HUMAN PRELIMINARY; PRT; 125 AA.
 AC Q6QBN8
 DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
 DT 05-JUL-2004, sequence version 1.
 DT 07-FEB-2006, entry version 8.
 DE T-cell immune regulator 1 transcript variant 4 (Fragment).
 GN Name=TCIRG1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC Tissue=Tonsil;
 RX PubMed=15809087; DOI=10.1016/j.bbrc.2005.03.065;
 RA Smirnova A.S., Morgun A., Shulzhenko N., Silva I.D.,
 RA Gerbase-Delima M.;
 RT "Identification of new alternative splice events in the TCIRG1 gene in
 RT different human tissues";
 RL Biochem. Biophys. Res. Commun. 330:943-949(2005).
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC
 DR EMBL; AY548967; AAS59834.1; -; mRNA.
 DR Ensembl; ENSG00000110719; Homo sapiens.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0015992; P:proton transport; IEA.
 DR InterPro; IPR002490; V_ATPase_sub116.
 DR PANTHER; PTHR11629; V_ATPase_sub116; 1.
 DR Pfam; PF01496; V_ATPase_I; 1.
 DR NON_TER 125 125
 SQ SEQUENCE 125 AA; 14272 MW; C7AA604E66F54444 CRC64;
 Query Match 7.1%; Score 58; DB 2; Length 125;
 Best Local Similarity 100.0%; Pred. No. 3.7e-50;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 57 RCELEKFTTFLQEVRRAGLVLPKGLPAPPDRDLRIQETTERLAQELRDVRGN 114
 DB 57 RCELEKFTTFLQEVRRAGLVLPKGLPAPPDRDLRIQETTERLAQELRDVRGN 114
 RESULT 6

Q8TCH1_HUMAN
ID Q8TCH1_HUMAN PRELIMINARY; PRT; 61 AA.
AC Q8TCH1;
DT 01-JUN-2002, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2002, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE T-cell immune regulator 1 transcript variant 3 (Fragment).
DE Name=TCIRG1;
GN Name=TCIRG1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterine cervix;
RX PubMed=22698286; PubMed=12813892;
RA Shulzhenko N., Smirnova A.S., Morgun A., Gerbase-DeLima M.;
RT "Specificity of alternative splice form detection using RT-PCR with a
RT primer spanning the exon junction."
RL BioTechniques 34:1244-1249 (2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Uterine cervix;
RX PubMed=15809087; DOI=10.1016/j.bbrc.2005.03.065;
RA Smirnova A.S., Morgun A., Shulzhenko N., Silva I.D.,
RA Gerbase-DeLima M.;
RT "Identification of new alternative splice events in the TCIRG1 gene in
RT different human tissues."
RL Biochem. Biophys. Res. Commun. 330:943-949 (2005).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF497545; AAM18704.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR002490; V_ATPase_sub116.
DR PANTHER; PTHR11629; V_ATPase_sub116; 1.
DR Pfam; PF01496; V_ATPase_1; 1.
FT NON_TER 1
FT SEQUENCE 61 AA; 6805 MW; B5F91060BFB3D01 CRC64;

Query Match 5.6%; Score 46; DB 2; Length 61;
Best Local Similarity 100.0%; Pred. No. 4.2e-38;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 699 DDEEAEELVPSEVLHQAIHTIEFCLGCVSNTASYLRWLWALSLAHA 744
DB 1 DDEEAEELVPSEVLHQAIHTIEFCLGCVSNTASYLRWLWALSLAHA 46

RESULT 7
Q9CTA9_MOUSE
ID Q9CTA9_MOUSE PRELIMINARY; PRT; 490 AA.
AC Q9CTA9;
DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2001, sequence version 1.
DT 07-FEB-2006, entry version 22.
DE 18-day embryo whole body cDNA. RIKEN full-length enriched library.
DE clone:1110034K16 product:ATPASE, H+ TRANSPORTING, LYSOSOMAL I, full
DE insert sequence. (Fragment).
GN Name=rcirg1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Cluttenbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagiolini M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Huminecki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugitara K., Sultana R., Takenaka Y., Taki K.,
RA Tamajima K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Tesdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Akashira N.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Nimomiya N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563 (2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566 (2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.B.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Kraglov D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perte G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,

RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wellie C.,
RA Wilning L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sakaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saio T., Okazaki Y., Gojobori T., Bono H., Kaeukawa T., Saio R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glessi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [8]

RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBAJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AK004094; BAB231166.1; -; mRNA.
DR Ensembl; ENSMUSG0000001750; Mus musculus.
DR MGI; MGI:1350931; Tcirlg.
DR GO; GO:0016021; C:integral to membrane; RCA.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0008553; F:hydrogen-exporting ATPase activity, phospho. .; RCA.
DR GO; GO:0006818; P:hydrogen transport; RCA.
DR InterPro; IPR002490; V_ATPase_sub116.
DR PANTHER; PTHR11629; V_ATPase_sub116; 1.
DR Pfam; PF01496; V_ATPase_1; 1.
FT NON_TER 1
SQ SEQUENCE 490 AA; 54595 MW; 89687F036A47542A CRC64;
Query Match 4.9%; Score 40; DB 2; Length 490;
Best Local Similarity 100.0%; Pred. No. 3.7e-31; Mismatches 0; Indels 0; Gaps 0;
Matches 40; Conservative 0;
QY 783 TVALLVMEGLSFAFLHALRLHWVEFQNKFSYGTGYKLSPP 822
DB 444 TVALLVMEGLSFAFLHALRLHWVEFQNKFSYGTGYKLSPP 483
RESULT 8
QJHFS_MOUSE PRELIMINARY; PRT; 834 AA.
AC QJHFS_MOUSE
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-FEB-2006, entry version 18.
DE A3 subunit of vacuolar-adenosine triphosphatase (Osteoclast-specific
DE 116-kDa V-ATPase subunit).
GN Name=Tcirlg; Synonym=Oc116;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20187595; PubMed=10722719; DOI=10.1074/jbc.275.12.8760;
RA Toyomura T., Oka T., Yamaguchi C., Wada Y., Futai M.;
RT "Three subunit a isoforms of mouse vacuolar H+-ATPase. Preferential
RT expression of the a3 isoform during osteoclast differentiation.";
RL J. Biol. Chem. 275:8760-8765 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=129/Sv;
RX MEDLINE=20173225; PubMed=10709991; DOI=10.1016/S8756-3282(99)00278-1;
RA Scimeca J.-C., Franchi A., Trojani C., Parrinello H., Grosgeorge J.,
RA Robert C., Jaillon O., Poirier C., Gaudray P., Carle G.P.;
RT "The gene encoding the mouse homologue of the human osteoclast-
RT specific 116-kDa V-ATPase subunit bears a deletion in osteoclastotic
RT (oc/oc) mutants.";
RL Bone 26:207-213 (2000).
RN [3]
RP NUCLEOTIDE SEQUENCE.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AB022322; BAA33006.1; -; mRNA.
DR EMBL; AF188702; AAF37193.1; -; Genomic DNA.
DR Ensembl; ENSMUSG0000001750; Mus musculus.
DR MGI; MGI:1350931; Tcirlg.
DR GO; GO:0016021; C:integral to membrane; RCA.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0008553; F:hydrogen-exporting ATPase activity, phospho. .; RCA.
DR GO; GO:0006818; P:hydrogen transport; RCA.
DR InterPro; IPR002490; V_ATPase_sub116.

```
DR PANTHER: PTHR11629; V_ATPase_sub116; 1.
DR Pfam: PF01496; V_ATPase_I; 1.
SQ SEQUENCE 834 AA; 93460 MW; E95ECD70C26367C0 CRC64;

Query Match      4.9%; Score 40; DB 2; Length 834;
Best Local Similarity 100.0%; Pred. No. 5.8e-31;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 783 TVAILLWMEGLSAFLHALRLHWVEFQNKFGYSGTGKLSPPF 822
Db 788 TVAILLWMEGLSAFLHALRLHWVEFQNKFGYSGTGKLSPPF 827

RESULT 9
Q9JL12_MOUSE
ID Q9JL12_MOUSE PRELIMINARY; PRT; 834 AA.
AC Q9JL12;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Vacuolar proton-translocating ATPase 100 kDa subunit isoform a3.
GN Name=rcirg1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=20167151; PubMed=10702241; DOI=10.1074/jbc.275.10.6824;
RA Nishi T., Forgac M.;
RT "Molecular cloning and expression of three isoforms of the 100-kDa a
subunit of the mouse vacuolar proton-translocating ATPase.";
RL J. Biol. Chem. 275:6824-6830(2000).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL: AF218253; AAP59922.1; -; mRNA.
CC Ensembl: ENSMUSG0000001750; Mus musculus.
CC MGI: MGI:1350931; Tcirlg1.
CC GO: GO:0016021; C:integral to membrane; RCA.
CC GO: GO:0005886; C:plasma membrane; IDA.
CC GO: GO:0008553; F:hydrogen-exporting ATPase activity, phospho. .; RCA.
CC GO: GO:0006818; P:hydrogen transport; RCA.
CC InterPro: IPR002490; V_ATPase_sub116.
CC PANTHER: PTHR11629; V_ATPase_sub116; 1.
CC Pfam: PF01496; V_ATPase_I; 1.
CC SEQUENCE 834 AA; 93506 MW; EB8E42E45163004F CRC64;

Query Match      4.9%; Score 40; DB 2; Length 834;
Best Local Similarity 100.0%; Pred. No. 5.8e-31;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 783 TVAILLWMEGLSAFLHALRLHWVEFQNKFGYSGTGKLSPPF 822
Db 788 TVAILLWMEGLSAFLHALRLHWVEFQNKFGYSGTGKLSPPF 827

RESULT 10
Q9JW06_MOUSE
ID Q9JW06_MOUSE PRELIMINARY; PRT; 834 AA.
AC Q9JW06;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-MAR-2006, entry version 20.
DE T-cell, immune regulator 1.
GN Name=rcirg1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=20167151; PubMed=10702241; DOI=10.1074/jbc.275.10.6824;
RA Nishi T., Forgac M.;
RT "Molecular cloning and expression of three isoforms of the 100-kDa a
subunit of the mouse vacuolar proton-translocating ATPase.";
RL J. Biol. Chem. 275:6824-6830(2000).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL: AF218253; AAP59922.1; -; mRNA.
CC Ensembl: ENSMUSG0000001750; Mus musculus.
CC MGI: MGI:1350931; Tcirlg1.
CC GO: GO:0016021; C:integral to membrane; RCA.
CC GO: GO:0005886; C:plasma membrane; IDA.
CC GO: GO:0008553; F:hydrogen-exporting ATPase activity, phospho. .; RCA.
CC GO: GO:0006818; P:hydrogen transport; RCA.
CC InterPro: IPR002490; V_ATPase_sub116.
CC PANTHER: PTHR11629; V_ATPase_sub116; 1.
CC Pfam: PF01496; V_ATPase_I; 1.
CC SEQUENCE 834 AA; 93506 MW; EB8E42E45163004F CRC64;

Query Match      4.9%; Score 40; DB 2; Length 834;
Best Local Similarity 100.0%; Pred. No. 5.8e-31;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 783 TVAILLWMEGLSAFLHALRLHWVEFQNKFGYSGTGKLSPPF 822
Db 788 TVAILLWMEGLSAFLHALRLHWVEFQNKFGYSGTGKLSPPF 827

RESULT 11
Q2I6B0_RAT
ID Q2I6B0_RAT PRELIMINARY; PRT; 834 AA.
AC Q2I6B0;
DT 07-MAR-2006, integrated into UniProtKB/TrEMBL.
DT 07-MAR-2006, sequence version 1.
DT 07-MAR-2006, entry version 1.
DE V-H-ATPase subunit a3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary tumor. C3, and Mammary tumor.
RC WAP-Tag model. 5 months old;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
Rahs S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NMRI; TISSUE=Mammary tumor. WAP-Tag model. 5 months old;
RG NIH MGC Project;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RG NIH MGC Project;
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL: BC006761; AA06761.1; -; mRNA.
CC EMBL: BC085234; AA085234.1; -; mRNA.
CC Ensembl: ENSMUSG0000001750; Mus musculus.
CC MGI: MGI:1350931; Tcirlg1.
CC GO: GO:0016021; C:integral to membrane; RCA.
CC GO: GO:0005886; C:plasma membrane; IDA.
CC GO: GO:0008553; F:hydrogen-exporting ATPase activity, phospho. .; RCA.
CC GO: GO:0006818; P:hydrogen transport; RCA.
CC InterPro: IPR002490; V_ATPase_sub116.
CC PANTHER: PTHR11629; V_ATPase_sub116; 1.
CC Pfam: PF01496; V_ATPase_I; 1.
CC SEQUENCE 834 AA; 93448 MW; 87D2C37911EBF172 CRC64;

Query Match      4.9%; Score 40; DB 2; Length 834;
Best Local Similarity 100.0%; Pred. No. 5.8e-31;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 783 TVAILLWMEGLSAFLHALRLHWVEFQNKFGYSGTGKLSPPF 822
Db 788 TVAILLWMEGLSAFLHALRLHWVEFQNKFGYSGTGKLSPPF 827
```

OC Muroidea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=wisar;
RA Guyon S., Amar M., Foosier P., Morel N.;
RT "Neurons coexpress three different v-ATPase subunit a isoforms.";
RL Submitted (NOV-2005) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; DQ286426; ABB91445.1; -; mRNA.
SQ SEQUENCE 834 AA; 93168 MW; 2E0DCFF9EED3979B CRC64;

Query Match 4.9%; Score 40; DB 2; Length 834;
Best Local Similarity 100.0%; Pred. No. 5.8e-31;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 783 TVALLVMEGLSAPFLHALRHVWFQNKFSYGTGYKLSPF 822
|||||
DB 788 TVALLVMEGLSAPFLHALRHVWFQNKFSYGTGYKLSPF 827
|||||

RESULT 12
Q6P735 RAT PRELIMINARY; PRT; 420 AA.
AC Q6P735
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE Tcigr1 protein.
GN Name=Tcigr1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidea; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Prostate;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stepleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Prostate;
RG NIH MGC Project;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC061859; AAH61859.1; -; mRNA.
DR Ensembl; ENSRNOG00000017220; Rattus norvegicus.

DR RGD; 735136; Tcigr1.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR002490; V_ATPase_sub116.
DR PANTHER; PTHR11629; V_ATPase_sub116; 1.
DR Pfam; PF01496; V_ATPase_I; 1.
SQ SEQUENCE 420 AA; 46956 MW; 1D3E094160F1ABD0 CRC64;

Query Match 4.6%; Score 38; DB 2; Length 420;
Best Local Similarity 100.0%; Pred. No. 3.6e-29;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 EVNPAPYTIITPPFLPAVMFGDVGHGLLMFLPALAMVL 422
|||||
DB 247 EVNPAPYTIITPPFLPAVMFGDVGHGLLMFLPALAMVL 284
|||||

RESULT 13
Q6QBN7 HUMAN PRELIMINARY; PRT; 45 AA.
AC Q6QBN7
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DT 07-FEB-2006, entry version 8.
DE T-cell immune regulator 1 transcript variant 5 (Fragment).
GN Name=TCIRG1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Spleen;
RX PubMed=15809087; DOI=10.1016/j.jbbr.2005.03.065;
RA Smirnova A.S., Morgun A., Shulzhenko N., Silva I.D.,
RA Gerbase-Delima M.;
RT "Identification of new alternative splice events in the TCIRG1 gene in
different human tissues.";
RL Biochem. Biophys. Res. Commun. 330:943-949 (2005).
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AY548968; AAS59835.1; -; mRNA.
DR Ensembl; ENSG00000110719; Homo sapiens.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR002490; V_ATPase_sub116.
DR PANTHER; PTHR11629; V_ATPase_sub116; 1.
DR Pfam; PF01496; V_ATPase_I; 1.
FT NON_TER 1
SQ SEQUENCE 45 AA; 5125 MW; 08E1883DECB81C70 CRC64;

Query Match 4.3%; Score 35; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 6.4e-27;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 453 FSIYTGFIYNECFRSATSIFFPSGWSVAAMANSQGW 487
|||||
DB 10 FSIYTGFIYNECFRSATSIFFPSGWSVAAMANSQGW 44
|||||

RESULT 14
Q7ZVM7 BRARE PRELIMINARY; PRT; 822 AA.
AC Q7ZVM7
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 14.
DE Zgc:55891.
GN ORFName=zgc:55891;
OS Brachydanio rerio (Zebrafish) (Danio rerio).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=AB; TISSUE=Whole body;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Sapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitling M., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=AB; TISSUE=Whole body;
 RA Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 CC EMBL; BC045484; AA45484.1; -; mRNA.
 DR Ensembl; ENSDARG0000035711; Danio rerio.
 DR ZFIN; ZDB-GENE-040426-2022; zgc:55891.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0015992; P:proton transport; IEA.
 DR InterPro; IPR002490; V_ATPase_sub116.
 DR PANTHER; PTHR11629; V_ATPase_sub116; 1.
 DR Pfam; PF01496; V_ATPase_I; 1.
 SQ SEQUENCE 822 AA; 93476 MW; 9432428966111755 CRC64;
 Query Match 4.1%; Score 34; DB 2; Length 822;
 Best Local Similarity 100.0%; Pred. No. 8.3e-25;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 786 ILLVMEGLSAFLHALRLHWVEFQNKFYSGTGYKL 819
 DB 776 ILLVMEGLSAFLHALRLHWVEFQNKFYSGTGYKL 809
 RESULT 15
 Q6PA83 XENLA
 ID Q6PA83 XENLA PRELIMINARY; PRT; 846 AA.
 AC Q6PA83;
 DT 05-JUL-2004, integrated into UniprotKB/TrEMBL.
 DT 05-JUL-2004, sequence version 1.
 DT 07-FEB-2006, entry version 15.
 DE MGC68661 protein.
 GN Name=MGC68661;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RN NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Sapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitling M., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RN NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative."
 RL Dev. Dyn. 225:384-391(2002).
 RN [3]
 RN NUCLEOTIDE SEQUENCE.
 RC TISSUE=Kidney;
 RA Klein S., Strausberg R.;
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
 CC Distributed under the Creative Commons Attribution-NoDerivs License
 CC -----
 CC EMBL; BC060417; AA60417.1; -; mRNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0015992; P:proton transport; IEA.
 DR InterPro; IPR002490; V_ATPase_sub116.
 DR PANTHER; PTHR11629; V_ATPase_sub116; 1.
 DR Pfam; PF01496; V_ATPase_I; 1.
 SQ SEQUENCE 846 AA; 97025 MW; DFCBABA953780759 CRC64;
 Query Match 3.9%; Score 32; DB 2; Length 846;
 Best Local Similarity 100.0%; Pred. No. 9.7e-23;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 783 TVAILVMEGLSAFLHALRLHWVEFQNKFYSG 814
 DB 795 TVAILVMEGLSAFLHALRLHWVEFQNKFYSG 826

Search completed: June 29, 2006, 13:22:03
 Job time : 309 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2006, 13:22:20 ; Search time 52 Seconds
(without alignments)
1383.657 Million cell updates/sec

Title: US-10-783-519-2
Perfect score: 822
Sequence: 1 MGSMPFSEVALVOLFLPTA.....HWVEFQNKFGTGYKLSPF 822

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 650591 seqs, 87530628 residues

Word size : 1

Total number of hits satisfying chosen parameters: 649417

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:*
1: /EMC Celerra_SID33/prodata/2/iaa/5 COMB.pep.*
2: /EMC Celerra_SID33/prodata/2/iaa/6 COMB.pep.*
3: /EMC Celerra_SID33/prodata/2/iaa/7 COMB.pep.*
4: /EMC Celerra_SID33/prodata/2/iaa/H COMB.pep.*
5: /EMC Celerra_SID33/prodata/2/iaa/PCUTUS COMB.pep.*
6: /EMC Celerra_SID33/prodata/2/iaa/RE COMB.pep.*
7: /EMC Celerra_SID33/prodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------|
| 1 | 822 | 100.0 | 822 | 2 | US-08-684-932A-38 |
| 2 | 822 | 100.0 | 822 | 2 | US-09-618-304B-2 |
| 3 | 545 | 66.3 | 849 | 2 | US-09-349-016-8846 |
| 4 | 545 | 66.3 | 849 | 2 | US-09-349-016-8847 |
| 5 | 545 | 66.3 | 853 | 2 | US-09-949-016-8275 |
| 6 | 545 | 66.3 | 853 | 2 | US-09-949-016-8276 |
| 7 | 433 | 52.7 | 750 | 2 | US-09-949-016-7201 |
| 8 | 433 | 52.7 | 750 | 2 | US-09-949-016-7202 |
| 9 | 33 | 4.0 | 33 | 2 | US-09-618-304B-9 |
| 10 | 32 | 3.9 | 32 | 2 | US-09-618-304B-4 |
| 11 | 30 | 3.6 | 30 | 2 | US-09-618-304B-7 |
| 12 | 26 | 3.2 | 26 | 2 | US-09-618-304B-5 |
| 13 | 24 | 2.9 | 513 | 2 | US-10-094-749-2211 |
| 14 | 24 | 2.9 | 847 | 2 | US-09-618-304B-3 |
| 15 | 23 | 2.8 | 23 | 2 | US-09-618-304B-8 |
| 16 | 22 | 2.7 | 155 | 2 | US-09-270-767-31642 |
| 17 | 22 | 2.7 | 155 | 2 | US-09-270-767-46859 |
| 18 | 22 | 2.7 | 257 | 2 | US-09-270-767-33142 |
| 19 | 22 | 2.7 | 257 | 2 | US-09-270-767-48359 |
| 20 | 21 | 2.6 | 21 | 2 | US-09-618-304B-6 |
| 21 | 20 | 2.4 | 380 | 2 | US-09-270-767-44518 |
| 22 | 20 | 2.4 | 855 | 7 | 5196526-1 |
| 23 | 18 | 2.2 | 196 | 2 | US-09-248-796A-17445 |
| 24 | 18 | 2.2 | 386 | 2 | US-09-248-796A-17451 |
| 25 | 17 | 2.1 | 163 | 2 | US-09-270-767-43040 |
| 26 | 13 | 1.6 | 343 | 2 | US-09-270-767-45290 |

27 13 1.6 515 2 US-09-270-767-43227 Sequence 43227, A
28 9 1.1 268 2 US-09-252-991A-20242 Sequence 20242, A
29 9 1.1 719 2 US-09-328-352-6274 Sequence 6274, Ap
30 9 1.1 833 2 US-09-252-991A-25227 Sequence 25227, A
31 8 1.0 134 2 US-09-543-681A-8295 Sequence 8295, Ap
32 8 1.0 187 1 US-08-441-629-11 Sequence 11, Appl
33 8 1.0 187 2 US-08-776-207-11 Sequence 11, Appl
34 8 1.0 187 2 US-09-507-773-11 Sequence 11, Appl
35 8 1.0 187 2 US-10-016-447-11 Sequence 11, Appl
36 8 1.0 187 5 PCT-US95-09172-11 Sequence 11, Appl
37 8 1.0 232 2 US-09-270-767-31821 Sequence 31821, A
38 8 1.0 295 2 US-09-248-796A-17556 Sequence 17556, A
39 8 1.0 299 2 US-10-094-749-1848 Sequence 1848, Ap
40 8 1.0 328 2 US-09-252-991A-27822 Sequence 27822, A
41 8 1.0 339 2 US-09-252-991A-19249 Sequence 19249, A
42 8 1.0 360 2 US-09-107-532A-4523 Sequence 4523, Ap
43 8 1.0 371 2 US-09-252-991A-23484 Sequence 23484, A
44 8 1.0 391 2 US-09-489-039A-7804 Sequence 7804, Ap
45 8 1.0 399 2 US-09-252-991A-19103 Sequence 19103, A

ALIGNMENTS

RESULT 1
US-08-684-932A-38
; Sequence 38, Application US/08684932A
; Patent No. 6403304
; GENERAL INFORMATION:
; APPLICANT: Stashenko, Philip
; APPLICANT: Li, Yi-Ping
; APPLICANT: Wuchterpfennig, Anne L.
; TITLE OF INVENTION: HUMAN OSTEOCLAST-SPECIFIC AND -RELATED
; TITLE OF INVENTION: DNA SEQUENCES
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/684,932A
; FILING DATE: 19-JUL-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: FDC92-02FM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-684-932A-38

Query Match 100.0%; Score 822; DB 2; Length 822;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 822; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGSMPFSEVALVOLFLPTAAYTCVSRIGELGVSRFDLNASVSAPQRFVVDVWRCEE 60
Db 1 MGSMPFSEVALVOLFLPTAAYTCVSRIGELGVSRFDLNASVSAPQRFVVDVWRCEE 60

| | | | |
|----|-----|---|-----|
| Qy | 61 | LEKTTFLQEEVRRAGLVLPPPKGBELPAPPPRDLLIRIQEETERLQAELRDVGNQOALPA | 120 |
| Db | 61 | LEKTTFLQEEVRRAGLVLPPPKGBELPAPPPRDLLIRIQEETERLQAELRDVGNQOALPA | 120 |
| Qy | 121 | QLHQQLHAAVLRQGHQPOLAAAHDTGASERTPLQAPGGPHQDRLVNFPVAGAVEPHKAP | 180 |
| Db | 121 | QLHQQLHAAVLRQGHQPOLAAAHDTGASERTPLQAPGGPHQDRLVNFPVAGAVEPHKAP | 180 |
| Qy | 181 | ALERLLWRACRGFLIASPRELEQPIUEHPVTGSPATWMTFLISYWGEOIGQKIRKITDCPH | 240 |
| Db | 181 | ALERLLWRACRGFLIASPRELEQPIUEHPVTGSPATWMTFLISYWGEOIGQKIRKITDCPH | 240 |
| Qy | 241 | CHVPFPLQOEBEARLGALQOLQOOSQELQEVGETERFISOVLGRVLQLLPPQOVQVHKM | 300 |
| Db | 241 | CHVPFPLQOEBEARLGALQOLQOOSQELQEVGETERFISOVLGRVLQLLPPQOVQVHKM | 300 |
| Qy | 301 | AVYLALNQCSVSTTHKCLIAEAWCWSRDLPALQEARLDRSSMBEGVSVAHRIPCRDMPT | 360 |
| Db | 301 | AVYLALNQCSVSTTHKCLIAEAWCWSRDLPALQEARLDRSSMBEGVSVAHRIPCRDMPT | 360 |
| Qy | 361 | LIRNRRPTASQGIVDRYGVGRYQEVNPAPIITTFPFLFAVMFGDVGHLLMFLPALAM | 420 |
| Db | 361 | LIRNRRPTASQGIVDRYGVGRYQEVNPAPIITTFPFLFAVMFGDVGHLLMFLPALAM | 420 |
| Qy | 421 | VLAENRPAVKAAQNEIWOTFPRGRYLLLMGLFSIYTGFIYNECSRATSIIPSGWSVAA | 480 |
| Db | 421 | VLAENRPAVKAAQNEIWOTFPRGRYLLLMGLFSIYTGFIYNECSRATSIIPSGWSVAA | 480 |
| Qy | 481 | MANQSGWSDAFLAQHTMLTLDPNVTGVFLGPYPFGIDPIWLSAANHSLPNSFKMKMSVI | 540 |
| Db | 481 | MANQSGWSDAFLAQHTMLTLDPNVTGVFLGPYPFGIDPIWLSAANHSLPNSFKMKMSVI | 540 |
| Qy | 541 | LGVVHMAFGVVLGVFNHVFQGRHRLLETPELTFLGLFGYLVFLVIYKWLCVWAARA | 600 |
| Db | 541 | LGVVHMAFGVVLGVFNHVFQGRHRLLETPELTFLGLFGYLVFLVIYKWLCVWAARA | 600 |
| Qy | 601 | ASPSILIHFINMFLFSHSPSNRLLYPROEVQATLVVLALAMVPIILLGTPLHLHRRHR | 660 |
| Db | 601 | ASPSILIHFINMFLFSHSPSNRLLYPROEVQATLVVLALAMVPIILLGTPLHLHRRHR | 660 |
| Qy | 661 | RLRRPADRQENKAGLLDLDPDASVNGWSDBEEKAGGLDDEBEABLVPSEVLHQAIHTI | 720 |
| Db | 661 | RLRRPADRQENKAGLLDLDPDASVNGWSDBEEKAGGLDDEBEABLVPSEVLHQAIHTI | 720 |
| Qy | 721 | EFCLGCVSNTASYLRMLWALSIAHAQSEVLWAMVMRIIGLIGREGVGAIVLVPFPAFA | 780 |
| Db | 721 | EFCLGCVSNTASYLRMLWALSIAHAQSEVLWAMVMRIIGLIGREGVGAIVLVPFPAFA | 780 |
| Qy | 781 | VMTVAIILLVMEGLSAFTHALRLHWVEFQNKPYSGTGYKLSPPF | 822 |
| Db | 781 | VMTVAIILLVMEGLSAFTHALRLHWVEFQNKPYSGTGYKLSPPF | 822 |

| | | | | | | | | | | | | |
|-----------------------|-----|-----------------|-------------------|------------------|-------------|--------------|-----------|------------|----------|---------|--------|-----|
| US-09-618-304B-2 | | | | | | | | | | | | |
| Query Match | | 100.0%; | Score 822; | DB 2; | Length 822; | | | | | | | |
| Best Local Similarity | | 100.0%; | Pred. No. 0; | | | | | | | | | |
| Matches 822; | | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; | | | | | | | |
| Qy | 1 | MGSMFRSEEV | ALVOLFLPTAAATCVSR | LGELGVFRDLNASVSA | FORFVV | DVNRCEE | 60 | | | | | |
| Db | 1 | MGSMFRSEEV | ALVOLFLPTAAATCVSR | LGELGVFRDLNASVSA | FORFVV | DVNRCEE | 60 | | | | | |
| Qy | 61 | LEKTTFTLQ | BEVRAGLVLP | PPKGRLPAPP | PRDLLRIQ | ETERLAQELDRV | GNQQALRA | 120 | | | | |
| Db | 61 | LEKTTFTLQ | BEVRAGLVLP | PPKGRLPAPP | PRDLLRIQ | ETERLAQELDRV | GNQQALRA | 120 | | | | |
| Qy | 121 | QLHQLOLHA | AVLRQGH | EPOLAAAH | TGASERTFL | QAPGGPHODLR | VNFVAGAVE | PHKAP | 180 | | | |
| Db | 121 | QLHQLOLHA | AVLRQGH | EPOLAAAH | TGASERTFL | QAPGGPHODLR | VNFVAGAVE | PHKAP | 180 | | | |
| Qy | 181 | ALERLLWRA | CRGFLLIAS | FRELEQ | PLEHPVTG | EPATWMTFLIS | YWG | EQIGOKIRKI | TD | 240 | | |
| Db | 181 | ALERLLWRA | CRGFLLIAS | FRELEQ | PLEHPVTG | EPATWMTFLIS | YWG | EQIGOKIRKI | TD | 240 | | |
| Qy | 241 | CHVPFFLQ | BEARIGALQ | QQSQ | SELQ | EVLG | STERFL | SOVLGRV | LQLPPG | QV | 300 | |
| Db | 241 | CHVPFFLQ | BEARIGALQ | QQSQ | SELQ | EVLG | STERFL | SOVLGRV | LQLPPG | QV | 300 | |
| Qy | 301 | AVYLALNQC | SVSTTHK | CLIAEAC | SVRDLPAL | QEARLDS | SM | EGVSAVA | HRIPC | RM | 360 | |
| Db | 301 | AVYLALNQC | SVSTTHK | CLIAEAC | SVRDLPAL | QEARLDS | SM | EGVSAVA | HRIPC | RM | 360 | |
| Qy | 361 | LIRTNRTAS | FOGIVDRY | GVGRYQ | EVNP | APYTIITP | PF | FAVMFG | VGHGLL | MF | 420 | |
| Db | 361 | LIRTNRTAS | FOGIVDRY | GVGRYQ | EVNP | APYTIITP | PF | FAVMFG | VGHGLL | MF | 420 | |
| Qy | 421 | VLAENRPA | VAAQNEI | WQTFPR | GRYLLLMGL | PSIYTG | FIYNEC | PSRATSIP | PSG | SVAA | 480 | |
| Db | 421 | VLAENRPA | VAAQNEI | WQTFPR | GRYLLLMGL | PSIYTG | FIYNEC | PSRATSIP | PSG | SVAA | 480 | |
| Qy | 481 | MANQSG | SDAFLAQHT | MTLTD | PNVTGV | FLGPYP | FGIDP | PIWS | LAANHLS | FLNS | FKMK | 540 |
| Db | 481 | MANQSG | SDAFLAQHT | MTLTD | PNVTGV | FLGPYP | FGIDP | PIWS | LAANHLS | FLNS | FKMK | 540 |
| Qy | 541 | LGVVHMA | F | GVVLGV | PNHVHQ | BQRHLL | LETLPE | L | FLGLFG | VLFLVIY | KWLC | 600 |
| Db | 541 | LGVVHMA | F | GVVLGV | PNHVHQ | BQRHLL | LETLPE | L | FLGLFG | VLFLVIY | KWLC | 600 |
| Qy | 601 | ASPSIL | IHF | NMFLSH | SPSNRL | LYPRQ | EVQATLV | V | LALAMVPI | LLGTPL | HL | 660 |
| Db | 601 | ASPSIL | IHF | NMFLSH | SPSNRL | LYPRQ | EVQATLV | V | LALAMVPI | LLGTPL | HL | 660 |
| Qy | 661 | RLRRRPA | DROENKAG | L | LDLPDAS | VNG | SWG | SDEEKAG | GLDDEE | BAELVP | BEVLM | 720 |
| Db | 661 | RLRRRPA | DROENKAG | L | LDLPDAS | VNG | SWG | SDEEKAG | GLDDEE | BAELVP | BEVLM | 720 |
| Qy | 721 | EFCLG | CVSNTASY | LRLWALS | LAHQ | SEVLW | AMVMRIG | L | GLG | REVGVAA | VVLVPI | 780 |
| Db | 721 | EFCLG | CVSNTASY | LRLWALS | LAHQ | SEVLW | AMVMRIG | L | GLG | REVGVAA | VVLVPI | 780 |
| Qy | 781 | VMTVAI | LLVME | LSAFL | HALRL | HWBF | ONKFY | S | GTGYK | LS | SPF | 822 |
| Db | 781 | VMTVAI | LLVME | LSAFL | HALRL | HWBF | ONKFY | S | GTGYK | LS | SPF | 822 |

RESULT 3
US-09-949-016-8846
; Sequence 8846, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016

```

; CURRENT FILING DATE: 2000-04-14
; PRIORITY APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIORITY APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIORITY APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8846
; LENGTH: 849
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8846

```

| Query Match | 66.3% | Score 545; | DB 2; | Length 849; | | | | | | | | | | | |
|-----------------------|-----------------|---------------|-----------|-------------|-------|-------|------|-----|-------|-------|-----|-----|-----|-----|-----|
| Best Local Similarity | 99.7% | Pred. No. 0; | | | | | | | | | | | | | |
| Matches 765; | Conservative 0; | Mismatches 1; | Indels 1; | Gaps 1; | | | | | | | | | | | |
| QY | 57 | RCELEKTF | FLQOEVR | RAGLVLP | PPKGR | LPAPP | PRDL | RIQ | ESTER | LAQEL | RDL | VR | GNQ | 116 | |
| DB | 76 | RCELEKTF | FLQOEVR | RAGLVLP | PPKGR | LPAPP | PRDL | RIQ | ESTER | LAQEL | RDL | VR | GNQ | 135 | |
| QY | 117 | ALRAQL | HQLHQA | VLUR | QGH | EPOL | AAHT | DG | AS | ERT | FL | QA | PG | PHD | 176 |
| DB | 136 | ALRAQL | HQLHQA | VLUR | QGH | EPOL | AAHT | DG | AS | ERT | FL | QA | PG | PHD | 195 |
| QY | 177 | HKAPAL | ERLL | LLWR | AC | RG | FL | IA | S | PRE | LE | OP | L | HP | 236 |
| DB | 196 | HKAPAL | ERLL | LLWR | AC | RG | FL | IA | S | PRE | LE | OP | L | HP | 255 |
| QY | 237 | DCFCHV | FP | FLQOE | EAR | LGA | LQ | LQ | QO | S | QEL | QEL | QEL | QEL | 296 |
| DB | 256 | DCFCHV | FP | FLQOE | EAR | LGA | LQ | LQ | QO | S | QEL | QEL | QEL | QEL | 315 |
| QY | 297 | HKMKAV | YLAL | NC | SV | STH | TK | CL | IA | E | AW | C | S | VR | 356 |
| DB | 316 | HKMKAV | YLAL | NC | SV | STH | TK | CL | IA | E | AW | C | S | VR | 375 |
| QY | 357 | MPPTL | LR | TNR | FT | AS | FG | I | VD | R | Y | G | V | R | 416 |
| DB | 376 | MPPTL | LR | TNR | FT | AS | FG | I | VD | R | Y | G | V | R | 435 |
| QY | 417 | ALAMV | LA | EN | RP | AK | AA | QNE | I | W | OT | F | F | R | 476 |
| DB | 436 | ALAMV | LA | EN | RP | AK | AA | QNE | I | W | OT | F | F | R | 495 |
| QY | 477 | SVAAM | AN | Q | S | W | S | D | A | F | L | A | H | T | 536 |
| DB | 496 | SVAAM | AN | Q | S | W | S | D | A | F | L | A | H | T | 555 |
| QY | 537 | MSVIL | G | V | H | W | H | A | F | G | V | L | F | G | 596 |
| DB | 556 | MSVIL | G | V | H | W | H | A | F | G | V | L | F | G | 615 |
| QY | 597 | AAAA | S | - | PS | L | I | H | F | I | N | M | F | S | 655 |
| DB | 616 | AAAA | S | A | P | S | I | L | I | H | F | I | N | M | 675 |
| QY | 656 | HRHRR | L | R | R | P | A | D | R | E | N | K | A | G | 715 |
| DB | 676 | HRHRR | L | R | R | P | A | D | R | E | N | K | A | G | 735 |
| QY | 716 | AIHTI | E | F | C | L | G | C | S | N | T | A | S | Y | 775 |
| DB | 736 | AIHTI | E | F | C | L | G | C | S | N | T | A | S | Y | 795 |
| QY | 776 | FAAFA | V | N | T | A | I | L | V | M | E | G | L | S | 822 |
| DB | 796 | FAAFA | V | N | T | A | I | L | V | M | E | G | L | S | 842 |

RESULT 4

```

US-09-949-016-8847
; Sequence 8847, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFECTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8847
; LENGTH: 849
; TYPE: FRT
; ORGANISM: Human
US-09-949-016-8847

```

| Query Match | 66.3%; | Score 545; | DB 2; | Length 849; |
|-----------------------|-----------------|-----------------------|---------------|---|
| Best Local Similarity | 99.7%; | Pred. No. 0; | | |
| Matches 765; | Conservative 0; | Mismatches 1; | Indels 1; | Gaps 1; |
| Qy | 57 | RCELEKTFLOEBEVRAGLVLP | PPKGRUPAPP | PRDLRLRQESTERLAQELRDVRGNOQ 116 |
| Db | 76 | RCELEKTFLOEBEVRAGLVLP | PPKGRUPAPP | PRDLRLRQESTERLAQELRDVRGNOQ 135 |
| Qy | 117 | ALRAQLHQLOLHAVALRQGH | EPOLAAAHATD | CASERTPLQAPGCPHODLRNVFVAGAVEP 176 |
| Db | 136 | ALRAQLHQLOLHAVALRQGH | EPOLAAAHATD | CASERTPLQAPGCPHODLRNVFVAGAVEP 195 |
| Qy | 177 | HKAPALERLLWRACRGFLI | ASFRELEQPLEHP | VTGEPATMTWTFIISYWGEOIGQKIRKIT 236 |
| Db | 196 | HKAPALERLLWRACRGFLI | ASFRELEQPLEHP | VTGEPATMTWTFIISYWGEOIGQKIRKIT 255 |
| Qy | 237 | DCFHCHVFPFLOEBEARL | QALQLOQOSOB | EOVLGETERPLSOVLGRVLOLLPPGQVOV 296 |
| Db | 256 | DCFHCHVFPFLOEBEARL | QALQLOQOSOB | EOVLGETERPLSOVLGRVLOLLPPGQVOV 315 |
| Qy | 297 | HKMKAVYALALNQCSVS | THKCLIAEAWCS | VRDLPALQEARLDSMEGVSVAHRIPCRD 356 |
| Db | 316 | HKMKAVYALALNQCSVS | THKCLIAEAWCS | VRDLPALQEARLDSMEGVSVAHRIPCRD 375 |
| Qy | 357 | MPPTLIRTNRTFASFG | QIVDRYGVGRY | QOEVNPAPYTIITFPFLFAVMFGDVGHGLLMFLF 416 |
| Db | 376 | MPPTLIRTNRTFASFG | QIVDAYGVGRY | QOEVNPAPYTIITFPFLFAVMFGDVGHGLLMFLF 435 |
| Qy | 417 | ALAMVLAENRPVAKAAQ | NEIWTQFFRG | RYLLMLGLFSIYTGFTYNECFSRATSIIPSGW 476 |
| Db | 436 | ALAMVLAENRPVAKAAQ | NEIWTQFFRG | RYLLMLGLFSIYTGFTYNECFSRATSIIPSGW 495 |
| Qy | 477 | SVAAMANQSGWDAFL | AQHTMLTLD | PNVTVGVFLGPYFGIDPTWLSLAANHLSFLNLSFMOK 536 |
| Db | 496 | SVAAMANQSGWDAFL | AQHTMLTLD | PNVTVGVFLGPYFGIDPTWLSLAANHLSFLNLSFMOK 555 |
| Qy | 537 | MSVILGVVHMAFGVL | GVFNHVFHQ | RHRLLELTPELTFLGLGFLYLVFLVIYKWLVCW 596 |
| Db | 556 | MSVILGVVHMAFGVL | GVFNHVFHQ | RHRLLELTPELTFLGLGFLYLVFLVIYKWLVCW 615 |
| Qy | 597 | AARAAS-PSLIIHFIN | FLSHSPSNRL | LIPREVVQATLVWLALAMVPILLGTPTLHLL 655 |
| Db | 616 | AARAASAPSLIHF | INMFLSHSPSN | RLLYIPREVVQATLVWLALAMVPILLGTPTLHLL 675 |
| Qy | 656 | HRHRRRLRRPADRQ | ENKAGLDDL | PDASVNGWSSDEEKAGGLDDEEAEIIPSEVLMLHQ 715 |
| Db | 676 | HRHRRRLRRPADRQ | ENKAGLDDL | PDASVNGWSSDEEKAGGLDDEEAEIIPSEVLMLHQ 735 |
| Qy | 716 | AIHTIEFCLGCVSNT | ASYLRILWALS | LAHAQLSEVLWAMVMVRIGLGLQREVGVAAVLVP 775 |

Db 736 AIIHTIEFCLGCVSNTASYLRWLSLAHAQLSEVLWAMVRIGLGLGREVGAAVVLVPI 795
QY 776 FAAPAVMTVAIILLVMEGLSAPFLHALRLHWVFQNKFSYGTGYKLSPP 822
Db 796 FAAPAVMTVAIILLVMEGLSAPFLHALRLHWVFQNKFSYGTGYKLSPP 842

RESULT 5

US-09-949-016-8275
; Sequence 8275, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8275
; LENGTH: 853
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8275

Query Match 66.3%; Score 545; DB 2; Length 853;

Best Local Similarity 99.7%; Pred. No. 0;
Matches 765; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 57 RCEELEKTFTFLOQEEVRRAGLVLPKGRLLPAPPPRDLRLIOETTERLAQELRDVRGNOQ 116
Db 80 RCEELEKTFTFLOQEEVRRAGLVLPKGRLLPAPPPRDLRLIOETTERLAQELRDVRGNOQ 139
QY 117 ALRAQLHQLHAALVLRQHEPQLAAHTDGAERTPLLOAPGGPHQDLRVNFVAGAVP 176
Db 140 ALRAQLHQLHAALVLRQHEPQLAAHTDGAERTPLLOAPGGPHQDLRVNFVAGAVP 199
QY 177 HKAPALERLLWRACRGFLIASPRELEQPLEHPVTGEPATMTMTFLISYWGEOIGQKIRKIT 236
Db 200 HKAPALERLLWRACRGFLIASPRELEQPLEHPVTGEPATMTMTFLISYWGEOIGQKIRKIT 259
QY 237 DCFHCHVFPFLOQEEARLGALQOQSOELQEVLTGERFELSQVLGRVLQLLPPQOVQV 296
Db 260 DCFHCHVFPFLOQEEARLGALQOQSOELQEVLTGERFELSQVLGRVLQLLPPQOVQV 319
QY 297 HKMAVYLALNCQSVSTTHKCLIAEAWCSVRDLPALQALRDSSMBEGVSAVAHRIPCRD 356
Db 320 HKMAVYLALNCQSVSTTHKCLIAEAWCSVRDLPALQALRDSSMBEGVSAVAHRIPCRD 379
QY 357 MPPTLIRNRTFTASQGIQVDAVGVRQYQEVNPAPTYITTFPFLFAVMFGDVGHGLMLFLF 416
Db 380 MPPTLIRNRTFTASQGIQVDAVGVRQYQEVNPAPTYITTFPFLFAVMFGDVGHGLMLFLF 439
QY 417 ALAMVLAENRPAKAAQNEIMQTFPRGRYLLLMGLFSIYTGFIYNECFSRATSIIPPSGW 476
Db 440 ALAMVLAENRPAKAAQNEIMQTFPRGRYLLLMGLFSIYTGFIYNECFSRATSIIPPSGW 499
QY 477 SVAAANQSGWSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWLSAANHLSFLNSFKMK 536
Db 500 SVAAANQSGWSDAFLAQHTMLTLDPNVTGVFLGYPFGIDPIWLSAANHLSFLNSFKMK 559
QY 537 MSVILGVVHMAFGVILGVFNHVFQGRHLLLETPELTFLGLFGYLVFLVIYKWLQVW 596
Db 560 MSVILGVVHMAFGVILGVFNHVFQGRHLLLETPELTFLGLFGYLVFLVIYKWLQVW 619

QY 597 AARAAS-PSILHIFNMFLFSHPNRLLYPRQEVVQATLVVLALAMVPTLLIGTPLHLL 655
Db 620 AARAASAPSILHIFNMFLFSHPNRLLYPRQEVVQATLVVLALAMVPTLLIGTPLHLL 679
QY 656 HRHRRRLRRPADROEENKAGLLDLPDASVNWSSDEEKAGGLDDEEAEELVSEVLMHQ 715
Db 680 HRHRRRLRRPADROEENKAGLLDLPDASVNWSSDEEKAGGLDDEEAEELVSEVLMHQ 739
QY 716 AIHTIEFCLGCVSNTASYLRWLSLAHAQLSEVLWAMVRIGLGLGREVGAAVVLVPI 775
Db 740 AIHTIEFCLGCVSNTASYLRWLSLAHAQLSEVLWAMVRIGLGLGREVGAAVVLVPI 799
QY 776 FAAPAVMTVAIILLVMEGLSAPFLHALRLHWVFQNKFSYGTGYKLSPP 822
Db 800 FAAPAVMTVAIILLVMEGLSAPFLHALRLHWVFQNKFSYGTGYKLSPP 846

RESULT 6

US-09-949-016-8276
; Sequence 8276, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8276
; LENGTH: 853
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8276

Query Match 66.3%; Score 545; DB 2; Length 853;

Best Local Similarity 99.7%; Pred. No. 0;
Matches 765; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 57 RCEELEKTFTFLOQEEVRRAGLVLPKGRLLPAPPPRDLRLIOETTERLAQELRDVRGNOQ 116
Db 80 RCEELEKTFTFLOQEEVRRAGLVLPKGRLLPAPPPRDLRLIOETTERLAQELRDVRGNOQ 139
QY 117 ALRAQLHQLHAALVLRQHEPQLAAHTDGAERTPLLOAPGGPHQDLRVNFVAGAVP 176
Db 140 ALRAQLHQLHAALVLRQHEPQLAAHTDGAERTPLLOAPGGPHQDLRVNFVAGAVP 199
QY 177 HKAPALERLLWRACRGFLIASPRELEQPLEHPVTGEPATMTMTFLISYWGEOIGQKIRKIT 236
Db 200 HKAPALERLLWRACRGFLIASPRELEQPLEHPVTGEPATMTMTFLISYWGEOIGQKIRKIT 259
QY 237 DCFHCHVFPFLOQEEARLGALQOQSOELQEVLTGERFELSQVLGRVLQLLPPQOVQV 296
Db 260 DCFHCHVFPFLOQEEARLGALQOQSOELQEVLTGERFELSQVLGRVLQLLPPQOVQV 319
QY 297 HKMAVYLALNCQSVSTTHKCLIAEAWCSVRDLPALQALRDSSMBEGVSAVAHRIPCRD 356
Db 320 HKMAVYLALNCQSVSTTHKCLIAEAWCSVRDLPALQALRDSSMBEGVSAVAHRIPCRD 379
QY 357 MPPTLIRNRTFTASQGIQVDAVGVRQYQEVNPAPTYITTFPFLFAVMFGDVGHGLMLFLF 416
Db 380 MPPTLIRNRTFTASQGIQVDAVGVRQYQEVNPAPTYITTFPFLFAVMFGDVGHGLMLFLF 439
QY 417 ALAMVLAENRPAKAAQNEIMQTFPRGRYLLLMGLFSIYTGFIYNECFSRATSIIPPSGW 476
Db 440 ALAMVLAENRPAKAAQNEIMQTFPRGRYLLLMGLFSIYTGFIYNECFSRATSIIPPSGW 499

| | | | |
|----|-----|--|-----|
| Qy | 477 | SVAAMANOSGWSDAFLAQHTMLTLPDPNVGTGVFLGPYPFGIDPISLAAHNSFLNSFKMK | 536 |
| Dd | 500 | SVAAMANOSGWSDAFLAQHTMLTLPDPNVGTGVFLGPYPFGIDPISLAAHNSFLNSFKMK | 559 |
| Qy | 537 | MSVILGVVHMAFGVVGLVPFNHVHFQQRHLLLETLPBELTLGLPGYLVFLVIYKWLVCWM | 596 |
| Dd | 560 | MSVILGVVHMAFGVVGLVPFNHVHFQQRHLLLETLPBELTLGLPGYLVFLVIYKWLVCWM | 619 |
| Qy | 597 | AARAAS - PSILIHFINMFLFSHPSPNRLLYPQREVVOATLVVLALAMWPIILLGTPLHL | 655 |
| Dd | 620 | AARAASPSII LHF INMFLFSHPSPNRLLYPQREVVOATLVVLALAMWPIILLGTPLHL | 679 |
| Qy | 656 | HRRRRLRRR PADROEENKAGLLDLPDASVNGSWSSDEEKAGGLDBEEAE LVPSEVLMHQ | 715 |
| Dd | 680 | HRRRRLRRR PADROEENKAGLLDLPDASVNGSWSSDEEKAGGLDBEEAE LVPSEVLMHQ | 739 |
| Qy | 716 | AIHTIEPCLCVSNTASYLR LWALS LAHAQA SEVLWAMVMRI GLGLGREVG VAAAVLVPI | 775 |
| Dd | 740 | AIHTIEFCLCVSNTASYLR LWALS LAHAQA SEVLWAMVMRI GLGLGREVG VAAAVLVPI | 799 |
| Qy | 776 | FAAFVMTVA ILLVMEGLSA FLHALRLHWBFONKFYS GTGYKLSPF | 822 |
| Dd | 800 | FAAFVMTVA ILLVMEGLSA FLHALRLHWBFONKFYS GTGYKLSPF | 846 |

RESULT 7
US-09-949-016-7201
; Sequence 7201, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7201
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7201

| | | | | |
|---|-----|--|-----|-----|
| Db | 329 | HGLLMFLFALAMVLAENRPAPVAKAONEIWQTFPRGRYLLLLMGLFSIYTGFIYNECF | SRA | 389 |
| Qy | 469 | TSIPSGSWAAMNQSGWSDAFIAQHTMLTLDPNVTGVLGYPFGIDPIWSLAANHLS | | 528 |
| Db | 389 | TSIPSGSWAAMNQSGWSDAFIAQHTMLTLDPNVTGVLGYPFGIDPIWSLAANHLS | | 448 |
| Qy | 529 | FLNSFKMKMSVILGVGMHAGCVLGVFNHVFQGRHLLLETLPELTLGLFGYLFLV | | 588 |
| Db | 449 | FLNSFKMKMSVILGVGMHAGCVLGVFNHVFQGRHLLLETLPELTLGLFGYLFLV | | 508 |
| Qy | 589 | IYKWLCVWAAAAA-PSILIHFINMFLFSHSPSNRLLYPROEVVQATLVVLAMVPI | | 647 |
| Db | 509 | IYKWLCVWAAAAA-SAPILIHFINMFLFSHSPSNRLLYPROEVVQATLVVLAMVPI | | 568 |
| Qy | 648 | LGTPLHLLHRRRLRRRPAQDRQENKAGLLDLPDASVNGSSDEEKAGGLDDEEAELV | | 707 |
| Db | 569 | LGTPLHLLHRRRLRRRPAQDRQENKAGLLDLPDASVNGSSDEEKAGGLDDEEAELV | | 628 |
| Qy | 708 | PSEVLHMQAHTTIEFCLGCVSNASYLRWLALSLAHQAQLSEVLWAMVRI | | 767 |
| Db | 629 | PSEVLHMQAHTTIEFCLGCVSNASYLRWLALSLAHQAQLSEVLWAMVRI | | 688 |
| Qy | 768 | AAVLVPIFAAFAMTVAILLVMGLSAPFLHALRLHWWVFQNKFSYGTGKLSPF | | 822 |
| Db | 689 | AAVLVPIFAAFAMTVAILLVMGLSAPFLHALRLHWWVFQNKFSYGTGKLSPF | | 743 |
| RESULT 8 | | | | |
| US-09-949-016-7202 | | | | |
| ; Sequence 7202, Application US/09949016 | | | | |
| ; Patent No. 6812339 | | | | |
| ; GENERAL INFORMATION: | | | | |
| ; APPLICANT: VENTER, J. Craig et al. | | | | |
| ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED | | | | |
| ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF | | | | |
| ; FILE REFERENCE: CL001307 | | | | |
| ; CURRENT APPLICATION NUMBER: US/09/949,016 | | | | |
| ; CURRENT FILING DATE: 2000-04-14 | | | | |
| ; PRIOR APPLICATION NUMBER: 60/241,755 | | | | |
| ; PRIOR FILING DATE: 2000-10-20 | | | | |
| ; PRIOR APPLICATION NUMBER: 60/237,768 | | | | |
| ; PRIOR FILING DATE: 2000-10-03 | | | | |
| ; PRIOR APPLICATION NUMBER: 60/231,498 | | | | |
| ; PRIOR FILING DATE: 2000-09-08 | | | | |
| ; NUMBER OF SEQ ID NOS: 207012 | | | | |
| ; SOFTWARE: Fast-Seq for Windows Version 4.0 | | | | |
| ; SEQ ID NO 7202 | | | | |
| ; LENGTH: 750 | | | | |
| ; TYPE: PRT | | | | |
| ; ORGANISM: Human | | | | |
| US-09-949-016-7202 | | | | |
| Query Match 52.7%; Score 433; DB 2; Length 750; | | | | |
| Best Local Similarity 99.7%; Pred. No. 0; | | | | |
| Matches 653; Conservative 0; Mismatches 1; Indels 1; Gaps 1; | | | | |
| Qy | 169 | FVAGAVEPHKAPALERLLWRACRGFLIASPRELEQPLEHPVTGEPATWMTFLISYNGEQI | | 228 |
| Db | 89 | FVAGAVEPHKAPALERLLWRACRGFLIASPRELEQPLEHPVTGEPATWMTFLISYNGEQI | | 148 |
| Qy | 229 | GQKIRKITDCHFCHVFPFFLQOEABRIGALQOQSQEQLQOEVLGETERFLS | | 288 |
| Db | 149 | GQKIRKITDCHFCHVFPFFLQOEABRIGALQOQSQEQLQOEVLGETERFLS | | 208 |
| Qy | 289 | LPFGQVQVHKMAVYLALNOCVSTTHKCLIAEAWCSVRDLPALQBALRDS | | 348 |
| Db | 209 | LPFGQVQVHKMAVYLALNOCVSTTHKCLIAEAWCSVRDLPALQBALRDS | | 268 |
| Qy | 349 | AHRIPCRDMPTLIIRNRTFTASFGQIVDVRGVGRYOEVNPAVTIITFPFLFAV | | 408 |
| Db | 269 | AHRIPCRDMPTLIIRNRTFTASFGQIVDVRGVGRYOEVNPAVTIITFPFLFAV | | 328 |
| Qy | 409 | HGLLMFLFALAMVLAENRPAPVAKAONEIWQTFPRGRYLLLLMGLFSIYTGFIYNECF | | 468 |

Db 329 HGLLAFALAVLAENRAVVAQAONEIQTFRRGRYLLLLMGLFSIYTGFIYNECFSA 388
QY 469 TSIPSGWSVAAMANSQSDAPLAQHTMLTLDPNVTGVFLGYPYPGIDPIWLSLAANHLS 528
Db 389 TSIPSGWSVAAMANSQSDAPLAQHTMLTLDPNVTGVFLGYPYPGIDPIWLSLAANHLS 448
QY 529 FLNSFKMKSVILGVVHMAFGVVLGVFNHVFQGRHRLLETLPELTFLGLFGYLVFLV 588
Db 449 FLNSFKMKSVILGVVHMAFGVVLGVFNHVFQGRHRLLETLPELTFLGLFGYLVFLV 508
QY 589 IYKMLCVWAARAAS-PSIILHINMELFSGHSPSNRLLYPRQEVQATLVVLALAMVPILL 647
Db 509 IYKMLCVWAARAASAPSIIILHINMELFSGHSPSNRLLYPRQEVQATLVVLALAMVPILL 568
QY 648 LGTPLHLHRRRLRRRRPADRQENKAGLLDLPDASVNGWSSDEKAGGLDDEEAELV 707
Db 569 LGTPLHLHRRRLRRRRPADRQENKAGLLDLPDASVNGWSSDEKAGGLDDEEAELV 628
QY 708 PSEVLHQAIHTIEFCLGCVSNATSYLRWLWLSLAHQISEVLWAMVMRIGLIGREVCV 767
Db 629 PSEVLHQAIHTIEFCLGCVSNATSYLRWLWLSLAHQISEVLWAMVMRIGLIGREVCV 688
QY 768 AAVLVPIFAAFVMTVAILLVMEGLSAFLHALRLHWHVEFQNKFSYGTGKLSPP 822
Db 689 AAVLVPIFAAFVMTVAILLVMEGLSAFLHALRLHWHVEFQNKFSYGTGKLSPP 743

RESULT 9

US-09-618-304B-9
; Sequence 9, Application US/09618304B
; Patent No. 6777537
; GENERAL INFORMATION:
; APPLICANT: Stashenko, Philip
; APPLICANT: Li, Yi-Ping
; TITLE OF INVENTION: Osteoclast Proton Pump Subunit
; FILE REFERENCE: 1564.1006-001
; CURRENT APPLICATION NUMBER: US/09/618,304B
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 08/605,378
; PRIOR FILING DATE: 1996-02-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-618-304B-9

Query Match 4.0%; Score 33; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.3e-21;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 764 EVGVAAVLVPIFAAFVMTVAILLVMEGLSAP 796
Db 1 EVGVAAVLVPIFAAFVMTVAILLVMEGLSAP 33

RESULT 10

US-09-618-304B-4
; Sequence 4, Application US/09618304B
; Patent No. 6777537
; GENERAL INFORMATION:
; APPLICANT: Stashenko, Philip
; APPLICANT: Li, Yi-Ping
; TITLE OF INVENTION: Osteoclast Proton Pump Subunit
; FILE REFERENCE: 1564.1006-001
; CURRENT APPLICATION NUMBER: US/09/618,304B
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 08/605,378
; PRIOR FILING DATE: 1996-02-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-618-304B-4

Query Match 3.9%; Score 32; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 9.9e-21;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 391 YTIITFPFLFAVMFGDVGHGLMFLFALAMVL 422
Db 1 YTIITFPFLFAVMFGDVGHGLMFLFALAMVL 32

RESULT 11

US-09-618-304B-7
; Sequence 7, Application US/09618304B
; Patent No. 6777537
; GENERAL INFORMATION:
; APPLICANT: Stashenko, Philip
; APPLICANT: Li, Yi-Ping
; TITLE OF INVENTION: Osteoclast Proton Pump Subunit
; FILE REFERENCE: 1564.1006-001
; CURRENT APPLICATION NUMBER: US/09/618,304B
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 08/605,378
; PRIOR FILING DATE: 1996-02-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-618-304B-7

Query Match 3.6%; Score 30; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.4e-19;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 571 LPFLTFLGLFGYLVFLVIYKWLVCVWAARA 600
Db 1 LPFLTFLGLFGYLVFLVIYKWLVCVWAARA 30

RESULT 12

US-09-618-304B-5
; Sequence 5, Application US/09618304B
; Patent No. 6777537
; GENERAL INFORMATION:
; APPLICANT: Stashenko, Philip
; APPLICANT: Li, Yi-Ping
; TITLE OF INVENTION: Osteoclast Proton Pump Subunit
; FILE REFERENCE: 1564.1006-001
; CURRENT APPLICATION NUMBER: US/09/618,304B
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 08/605,378
; PRIOR FILING DATE: 1996-02-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-618-304B-5

Query Match 3.2%; Score 26; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.6e-15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 QTFPRGRYLLLMGLFSIYTGFIYNE 463
Db 1 QTFPRGRYLLLMGLFSIYTGFIYNE 26

RESULT 13
US-10-094-749-2211
; Sequence 2211, Application US/10094749
; Patent No. 6979557
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2211
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2211

Query Match 2.9%; Score 24; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.3e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 728 SNTASYRLRLWALSIAHAQLSEVLW 751
Db 409 SNTASYRLRLWALSIAHAQLSEVLW 432

RESULT 14
US-09-618-304B-3
; Sequence 3, Application US/09618304B
; Patent No. 6777537
; GENERAL INFORMATION:
; APPLICANT: Stashenko, Philip
; APPLICANT: Li, Yi-Ping
; TITLE OF INVENTION: Osteoclast Proton Pump Subunit
; FILE REFERENCE: 1564.1006-001
; CURRENT APPLICATION NUMBER: US/09/618,304B
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 08/605,378
; PRIOR FILING DATE: 1996-02-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus sequence
; FEATURE:
; NAME/KEY: VARIANT

; LOCATION: (1)...(847)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-618-304B-3

Query Match 2.9%; Score 24; DB 2; Length 847;
Best Local Similarity 100.0%; Pred. No. 2.1e-12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 728 SNTASYRLRLWALSIAHAQLSEVLW 751
Db 741 SNTASYRLRLWALSIAHAQLSEVLW 764

RESULT 15
US-09-618-304B-8
; Sequence 8, Application US/09618304B
; Patent No. 6777537
; GENERAL INFORMATION:
; APPLICANT: Stashenko, Philip
; APPLICANT: Li, Yi-Ping
; TITLE OF INVENTION: Osteoclast Proton Pump Subunit
; FILE REFERENCE: 1564.1006-001
; CURRENT APPLICATION NUMBER: US/09/618,304B
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 08/605,378
; PRIOR FILING DATE: 1996-02-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-618-304B-8

Query Match 2.8%; Score 23; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 6.5e-13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 632 QATLVVLALAMVPILLGTPLHL 654
Db 1 QATLVVLALAMVPILLGTPLHL 23

Search completed: June 29, 2006, 13:23:53
Job time : 54 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
 Copyright (c) 1993 - 2006 Bioceleration Ltd.
 OM protein - protein search, using sw model
 Run on: June 29, 2006, 13:23:09 ; Search time 188 Seconds
 (without alignments)
 2025.333 Million cell updates/sec
 Title: US-10-783-519-2
 Perfect score: 822
 Sequence: 1 MGSMPRSEEVALLVQLFLPTA.....HWVEFQNKFSYGTGYKLSPF 822
 Scoring table: OLIGO
 Gapop 60.0 , Gapext 60.0
 Searched: 2097797 seqs, 463214858 residues

Word size : 1
 Total number of hits satisfying chosen parameters: 2096646

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA Main:
 1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
 2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
 3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
 4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pap.*
 5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pap.*
 6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 822 | 100.0 | 822 | 5 | US-10-783-519-2 |
| 2 | 385 | 46.8 | 614 | 4 | US-10-145-012-2 |
| 3 | 385 | 46.8 | 614 | 6 | US-11-126-866-2 |
| 4 | 385 | 46.8 | 614 | 6 | US-11-126-841A-2 |
| 5 | 325 | 39.5 | 614 | 4 | US-10-145-012-13 |
| 6 | 325 | 39.5 | 614 | 5 | US-10-489-725-5 |
| 7 | 325 | 39.5 | 614 | 6 | US-11-126-866-13 |
| 8 | 325 | 39.5 | 614 | 6 | US-11-126-841A-13 |
| 9 | 291 | 35.4 | 643 | 4 | US-10-264-049-2656 |
| 10 | 33 | 4.0 | 33 | 5 | US-10-783-519-9 |
| 11 | 32 | 3.9 | 32 | 5 | US-10-783-519-4 |
| 12 | 30 | 3.6 | 30 | 5 | US-10-783-519-7 |
| 13 | 29 | 3.5 | 29 | 5 | US-10-489-725-7 |
| 14 | 26 | 3.2 | 26 | 5 | US-10-783-519-5 |
| 15 | 24 | 2.9 | 24 | 5 | US-10-489-725-9 |
| 16 | 24 | 2.9 | 24 | 5 | US-10-094-749-2211 |
| 17 | 24 | 2.9 | 558 | 4 | US-10-450-763-39739 |
| 18 | 24 | 2.9 | 821 | 4 | US-10-017-161-2144 |
| 19 | 24 | 2.9 | 821 | 4 | US-10-292-798-1790 |
| 20 | 24 | 2.9 | 841 | 5 | US-10-874-706-20 |
| 21 | 24 | 2.9 | 847 | 5 | US-10-783-519-3 |
| 22 | 23 | 2.8 | 23 | 5 | US-10-783-519-8 |
| 23 | 22 | 2.7 | 834 | 6 | US-11-097-143-36642 |
| 24 | 22 | 2.7 | 834 | 6 | US-11-097-143-41499 |
| 25 | 22 | 2.7 | 834 | 6 | US-11-097-143-41700 |
| 26 | 22 | 2.7 | 844 | 6 | US-11-097-143-41496 |
| 27 | 22 | 2.7 | 873 | 4 | US-10-369-493-6070 |

| | | | | | | |
|----|----|-----|------|---|----------------------|--------------------|
| 28 | 22 | 2.7 | 1236 | 4 | US-10-369-493-6203 | Sequence 6203, Ap |
| 29 | 21 | 2.6 | 21 | 5 | US-10-783-519-6 | Sequence 6, Appli |
| 30 | 20 | 2.4 | 84 | 4 | US-10-424-599-258223 | Sequence 258223, A |
| 31 | 20 | 2.4 | 133 | 4 | US-10-767-701-40787 | Sequence 40787, A |
| 32 | 20 | 2.4 | 139 | 4 | US-10-425-115-334174 | Sequence 334174, A |
| 33 | 20 | 2.4 | 204 | 4 | US-10-425-114-67979 | Sequence 67979, A |
| 34 | 20 | 2.4 | 284 | 4 | US-10-425-115-334176 | Sequence 334176, A |
| 35 | 20 | 2.4 | 340 | 4 | US-10-424-599-258221 | Sequence 258221, A |
| 36 | 20 | 2.4 | 372 | 4 | US-10-424-599-255903 | Sequence 255903, A |
| 37 | 20 | 2.4 | 702 | 4 | US-10-389-566-1961 | Sequence 1961, Ap |
| 38 | 20 | 2.4 | 783 | 4 | US-10-437-963-152242 | Sequence 152242, A |
| 39 | 20 | 2.4 | 784 | 4 | US-10-437-963-196347 | Sequence 196347, A |
| 40 | 20 | 2.4 | 818 | 4 | US-10-495-446-32 | Sequence 32, Appli |
| 41 | 20 | 2.4 | 820 | 4 | US-10-437-963-196349 | Sequence 196349, A |
| 42 | 20 | 2.4 | 855 | 5 | US-10-714-995-14 | Sequence 14, Appli |
| 43 | 20 | 2.4 | 855 | 6 | US-11-097-143-2052 | Sequence 2052, Ap |
| 44 | 19 | 2.3 | 935 | 4 | US-10-369-493-5810 | Sequence 5810, Ap |
| 45 | 19 | 2.3 | 1030 | 4 | US-10-369-493-5811 | Sequence 5811, Ap |

ALIGNMENTS

RESULT 1

US-10-783-519-2
 ; Sequence 2, Application US/10783519
 ; Publication No. US200506448A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Stashenko, Philip
 ; APPLICANT: Li, Yi-Ping
 ; TITLE OF INVENTION: Osteoclast Proton Pump Subunit
 ; FILE REFERENCE: 1564.1006-001
 ; CURRENT APPLICATION NUMBER: US/10783,519
 ; CURRENT FILING DATE: 2004-02-20
 ; PRIOR APPLICATION NUMBER: US/09/618,304
 ; PRIOR FILING DATE: 2000-07-18
 ; PRIOR APPLICATION NUMBER: 08/605,378
 ; PRIOR FILING DATE: 1996-02-22
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 822
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; ORGANISM: Homo sapiens
 US-10-783-519-2

| | | | | | |
|-----------------------|-----|--|---------------|-----------|-------------|
| Query Match | | 100.0%; | Score 822; | DB 5; | Length 822; |
| Best Local Similarity | | 100.0%; | Pred. No. 0; | | |
| Matches 822; | | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 1 | MGSMPRSEEVALLVQLFLPTAAAYTCVSRGELGLVFRDLNASVSFAQRRFVVDVWRCCE | 60 | | |
| Db | 1 | MGSMPRSEEVALLVQLFLPTAAAYTCVSRGELGLVFRDLNASVSFAQRRFVVDVWRCCE | 60 | | |
| Qy | 61 | LEKTFTFLOEVRVAGLVLPKPKRPPAPPPRDLRIQETTERLAQLRDLVGRNQOALRA | 120 | | |
| Db | 61 | LEKTFTFLOEVRVAGLVLPKPKRPPAPPPRDLRIQETTERLAQLRDLVGRNQOALRA | 120 | | |
| Qy | 121 | QLHQLHQAVALVQLQGHEPOLAAHTDASERTPLQAPGGPHODLRVNFVAGAVEPHKAP | 180 | | |
| Db | 121 | QLHQLHQAVALVQLQGHEPOLAAHTDASERTPLQAPGGPHODLRVNFVAGAVEPHKAP | 180 | | |
| Qy | 181 | ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEGIQKIRKITDCFH | 240 | | |
| Db | 181 | ALERLLWRACRGFLIASFRELEQPLEHPVTGEPATWMTFLISYWGEGIQKIRKITDCFH | 240 | | |
| Qy | 241 | CHVFPFLOEVRVAGLVLPKPKRPPAPPPRDLRIQETTERLAQLRDLVGRNQOALRA | 300 | | |
| Db | 241 | CHVFPFLOEVRVAGLVLPKPKRPPAPPPRDLRIQETTERLAQLRDLVGRNQOALRA | 300 | | |
| Qy | 301 | AVYLAALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLSSMEEGVSAVAHRIPCRDMPPT | 360 | | |
| Db | 301 | AVYLAALNQCSVSTTHKCLIAEAWCSVRDLPALQEARLSSMEEGVSAVAHRIPCRDMPPT | 360 | | |

QY 361 LIRNRTASFOGIVDRYGVGRYQEVNPAPYITITPPFLFAVMFGDVGHLLMFLPALAM 420
DB 361 LIRNRTASFOGIVDRYGVGRYQEVNPAPYITITPPFLFAVMFGDVGHLLMFLPALAM 420
QY 421 VLAENRPVAKAAQNEITWQTFRGRYLLLLMGLFSIYTGFIYNECFSRATSIIPSGMSVAA 480
DB 421 VLAENRPVAKAAQNEITWQTFRGRYLLLLMGLFSIYTGFIYNECFSRATSIIPSGMSVAA 480
QY 481 MANQSGWSAPLAQHTMLTLDPNVTGVFLGPYPFGIDPIWSLAANHLSPNSFKMKSIVI 540
DB 481 MANQSGWSAPLAQHTMLTLDPNVTGVFLGPYPFGIDPIWSLAANHLSPNSFKMKSIVI 540
QY 541 LGVVMHAFGVVGVNFHVFHFGQRHRLLETLPELTLLGLFGLVFLVIYKMLCWAARA 600
DB 541 LGVVMHAFGVVGVNFHVFHFGQRHRLLETLPELTLLGLFGLVFLVIYKMLCWAARA 600
QY 601 ASPSILIHFINMFLPSHSPSNRLLYPRQEVQATLVVLAAMVPIILLGTPLHLHRRR 660
DB 601 ASPSILIHFINMFLPSHSPSNRLLYPRQEVQATLVVLAAMVPIILLGTPLHLHRRR 660
QY 661 RLRRPADROENKAGLLDLPDASVNGSSDEEKAGGLDDEBAELVPSEVLHQAIHTI 720
DB 661 RLRRPADROENKAGLLDLPDASVNGSSDEEKAGGLDDEBAELVPSEVLHQAIHTI 720
QY 721 EFCIGCVSNTASYLRWALSALHAQSEVLWAMVMRIGLGLREVGVAAVLPVIPAFA 780
DB 721 EFCIGCVSNTASYLRWALSALHAQSEVLWAMVMRIGLGLREVGVAAVLPVIPAFA 780
QY 781 VMTVAILLVMEGLSAPFLHALRLHWVEFQNKFSYSGTGKLSPP 822
DB 781 VMTVAILLVMEGLSAPFLHALRLHWVEFQNKFSYSGTGKLSPP 822

RESULT 2

US-10-145-012-2
; Sequence 2, Application US/10145012
; Publication No. US20030124614A1
; GENERAL INFORMATION:
; APPLICANT: UTU et al.
; TITLE OF INVENTION: T-CELL MEMBRANE PROTEIN (TIRC7), PEPTIDES AND ANTIBODIES DERIVED
; FILE REFERENCE: 4400-0105P
; CURRENT APPLICATION NUMBER: US/10/145,012
; CURRENT FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-145-012-2

Query Match 46.8%; Score 385; DB 4; Length 614;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 605; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 217 MTFLISYNGEIQGQIRKKTDCFHCHVFPFLOQEEARLGALQOQSQSELQEVLGETER 276
DB 1 MTFLISYNGEIQGQIRKKTDCFHCHVFPFLOQEEARLGALQOQSQSELQEVLGETER 60
QY 277 FLSQVLGRVLQLPPQGVQVHKMAVYLAALNOCVSTTHKCLIAEAWCSVRDLPALQEA 336
DB 61 FLSQVLGRVLQLPPQGVQVHKMAVYLAALNOCVSTTHKCLIAEAWCSVRDLPALQEA 120
QY 337 RDSSMEEGVSAVAHRIPCDMPPTLIRNRTFTASFGIVDRYGVGRYQEVNPAPYITIF 396
DB 121 RDSSMEEGVSAVAHRIPCDMPPTLIRNRTFTASFGIVDRYGVGRYQEVNPAPYITIF 180
QY 397 PFLPAMFGDVGHLLMFLPALAMVLAENRPVAKAAQNEIMQTFPRGRYLLLMGLFSY 456
DB 181 PFLPAMFGDVGHLLMFLPALAMVLAENRPVAKAAQNEIMQTFPRGRYLLLMGLFSY 240

QY 457 TGFYINECFSRATSIIPSGMSVAAMANQSGWSAPFLAQHTMLTLDPNVTGVFLGPYPFGI 516
DB 241 TGFYINECFSRATSIIPSGMSVAAMANQSGWSAPFLAQHTMLTLDPNVTGVFLGPYPFGI 300
QY 517 DPINSLAANHLSPNSFKMKSIVIILGVVHMAFGVVLGVNFHVFHFGQRHRLLETLPELTF 576
DB 301 DPINSLAANHLSPNSFKMKSIVIILGVVHMAFGVVLGVNFHVFHFGQRHRLLETLPELTF 360
QY 577 LLGLFGYLVFLVIYKMLCWAARAAS-PSILIHFINMFLPSHSPSNRLLYPRQEVQATL 635
DB 361 LLGLFGYLVFLVIYKMLCWAARAASPSILIHFINMFLPSHSPSNRLLYPRQEVQATL 420
QY 636 VVLAAMVPIILLGTPLHLHRRRLRRPADROENKAGLLDLPDASVNGSSDEEKA 695
DB 421 VVLAAMVPIILLGTPLHLHRRRLRRPADROENKAGLLDLPDASVNGSSDEEKA 480
QY 696 GGLDDEBAELVPSEVLHQAIHTIEFCIGCVSNTASYLRWALSALHAQSEVLWAMVM 755
DB 481 GGLDDEBAELVPSEVLHQAIHTIEFCIGCVSNTASYLRWALSALHAQSEVLWAMVM 540
QY 756 RIGLGLREVGVAAVLPVIPAFAVMTVAILLVMEGLSAPFLHALRLHWVEFQNKFSYSGT 815
DB 541 RIGLGLREVGVAAVLPVIPAFAVMTVAILLVMEGLSAPFLHALRLHWVEFQNKFSYSGT 600
QY 816 GYKLSPP 822
DB 601 GYKLSPP 607

RESULT 3

US-11-126-866-2
; Sequence 2, Application US/11126866
; Publication No. US20050220789A1
; GENERAL INFORMATION:
; APPLICANT: UTU, et al., NALAN
; TITLE OF INVENTION: METHODS OF TREATMENT USING AN ANTIBODY THAT SPECIFICALLY BINDS
; FILE REFERENCE: 1472/71099-ZA/JPM/AG
; CURRENT APPLICATION NUMBER: US/11/126,866
; CURRENT FILING DATE: 2005-05-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 2
; LENGTH: 614
; TYPE: PRT
; ORGANISM: human
US-11-126-866-2

Query Match 46.8%; Score 385; DB 6; Length 614;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 605; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 217 MTFLISYNGEIQGQIRKKTDCFHCHVFPFLOQEEARLGALQOQSQSELQEVLGETER 276
DB 1 MTFLISYNGEIQGQIRKKTDCFHCHVFPFLOQEEARLGALQOQSQSELQEVLGETER 60
QY 277 FLSQVLGRVLQLPPQGVQVHKMAVYLAALNOCVSTTHKCLIAEAWCSVRDLPALQEA 336
DB 61 FLSQVLGRVLQLPPQGVQVHKMAVYLAALNOCVSTTHKCLIAEAWCSVRDLPALQEA 120
QY 337 RDSSMEEGVSAVAHRIPCDMPPTLIRNRTFTASFGIVDRYGVGRYQEVNPAPYITIF 396
DB 121 RDSSMEEGVSAVAHRIPCDMPPTLIRNRTFTASFGIVDAVGGRYQEVNPAPYITIF 180
QY 397 PFLPAMFGDVGHLLMFLPALAMVLAENRPVAKAAQNEIMQTFPRGRYLLLMGLFSY 456
DB 181 PFLPAMFGDVGHLLMFLPALAMVLAENRPVAKAAQNEIMQTFPRGRYLLLMGLFSY 240
QY 457 TGFYINECFSRATSIIPSGMSVAAMANQSGWSAPFLAQHTMLTLDPNVTGVFLGPYPFGI 516
DB 241 TGFYINECFSRATSIIPSGMSVAAMANQSGWSAPFLAQHTMLTLDPNVTGVFLGPYPFGI 300
QY 517 DPINSLAANHLSPNSFKMKSIVIILGVVHMAFGVVLGVNFHVFHFGQRHRLLETLPELTF 576

Db 301 DPTWLSLAANHLSPNSFKMNSVILGVVHMAFGVVLGVFNHVFQGRHRLLETLPELT 360
QY 577 LLGLFGYLVFLVIYKMLCVWAARAAS-PSILIHFINMFLFSHSPSNRLLYPROEVQATL 635
Db 361 LLGLFGYLVFLVIYKMLCVWAARAASAPSIILIHFINMFLFSHSPSNRLLYPROEVQATL 420
QY 636 VVLALAMVPILLGLTPTLHLLHRRRLRRPADQENKAGLLDLPDASVNGWSDEKA 695
Db 421 VVLALAMVPILLGLTPTLHLLHRRRLRRPADQENKAGLLDLPDASVNGWSDEKA 480
QY 696 GGLDDEEAEELVPSEVLVLMHQAIIHTIEFCLGCVSNSTASYLRWALSALAHQSEVLWAMVM 755
Db 481 GGLDDEEAEELVPSEVLVLMHQAIIHTIEFCLGCVSNSTASYLRWALSALAHQSEVLWAMVM 540
QY 756 RIGLGLGREVGVAAVLVPIFAAFVAVMTVAILLVMEGLSAPFLHALRLHWHVEFQNKFSYGT 815
Db 541 RIGLGLGREVGVAAVLVPIFAAFVAVMTVAILLVMEGLSAPFLHALRLHWHVEFQNKFSYGT 600
QY 816 GYKLSPP 822
Db 601 GYKLSPP 607

RESULT 4

US-11-126-841A-2
; Sequence 2, Application US/11126841A
; Publication No. US20050271659A1

GENERAL INFORMATION:

; APPLICANT: UTKU, et al., NALAN
; TITLE OF INVENTION: METHODS OF TREATMENT USING AN ANTIBODY THAT SPECIFICALLY BINDS TO
; TITLE OF INVENTION: AN EXTRACELLULAR DOMAIN OF A TIRC7 MEMBRANE PROTEIN
; FILE REFERENCE: 1472/11099-ZB/JPW/AG
; CURRENT APPLICATION NUMBER: US/11/126.841A
; CURRENT FILING DATE: 2005-05-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 614
; TYPE: PRT
; ORGANISM: human
US-11-126-841A-2

Query Match 46.8%; Score 385; DB 6; Length 614;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 605; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 217 MTLISYWGSGIQGKIRKIDTCFCHVFPFLQOEAEALGALQQLQOQSBLQEVLGETER 276
Db 1 MTLISYWGSGIQGKIRKIDTCFCHVFPFLQOEAEALGALQQLQOQSBLQEVLGETER 60
QY 277 FLSQVLGRVLQLLPPGQVQVHKMAVYALNQCSTVTHKCLIAEAWCSVRDLPALQEAL 336
Db 61 FLSQVLGRVLQLLPPGQVQVHKMAVYALNQCSTVTHKCLIAEAWCSVRDLPALQEAL 120
QY 337 RDSMEEGSAVAHRIPCRDMPPTLIRTNFTASFGQIVDRYGVGRYQEVNPAITYITF 396
Db 121 RDSMEEGSAVAHRIPCRDMPPTLIRTNFTASFGQIVDRYGVGRYQEVNPAITYITF 180
QY 397 PFLFAMVFGDVGHGLMFLPALAMVLAENRPAVKAQNEIWOTFFRGYRLLILMLGFSY 456
Db 181 PFLFAMVFGDVGHGLMFLPALAMVLAENRPAVKAQNEIWOTFFRGYRLLILMLGFSY 240
QY 457 TGPYINCEFSRATSIFFSGWSVAAMANQSGWSDAFLAQHTMTLTDPNVTVGVFLGPYFPGI 516
Db 241 TGPYINCEFSRATSIFFSGWSVAAMANQSGWSDAFLAQHTMTLTDPNVTVGVFLGPYFPGI 300
QY 517 DPTWLSLAANHLSPNSFKMNSVILGVVHMAFGVVLGVFNHVFQGRHRLLETLPELT 576
Db 301 DPTWLSLAANHLSPNSFKMNSVILGVVHMAFGVVLGVFNHVFQGRHRLLETLPELT 360
QY 577 LLGLFGYLVFLVIYKMLCVWAARAAS-PSILIHFINMFLFSHSPSNRLLYPROEVQATL 635

Db 361 LLGLFGYLVFLVIYKMLCVWAARAASAPSIILIHFINMFLFSHSPSNRLLYPROEVQATL 420
QY 636 VVLALAMVPILLGLTPTLHLLHRRRLRRPADQENKAGLLDLPDASVNGWSDEKA 695
Db 421 VVLALAMVPILLGLTPTLHLLHRRRLRRPADQENKAGLLDLPDASVNGWSDEKA 480
QY 696 GGLDDEEAEELVPSEVLVLMHQAIIHTIEFCLGCVSNSTASYLRWALSALAHQSEVLWAMVM 755
Db 481 GGLDDEEAEELVPSEVLVLMHQAIIHTIEFCLGCVSNSTASYLRWALSALAHQSEVLWAMVM 540
QY 756 RIGLGLGREVGVAAVLVPIFAAFVAVMTVAILLVMEGLSAPFLHALRLHWHVEFQNKFSYGT 815
Db 541 RIGLGLGREVGVAAVLVPIFAAFVAVMTVAILLVMEGLSAPFLHALRLHWHVEFQNKFSYGT 600
QY 816 GYKLSPP 822
Db 601 GYKLSPP 607

RESULT 5

US-10-145-012-13
; Sequence 13, Application US/10145012
; Publication No. US20030124614A1

GENERAL INFORMATION:

; APPLICANT: UTKU et al.
; TITLE OF INVENTION: T-CELL MEMBRANE PROTEIN (TIRC7), PEPTIDES AND ANTIBODIES DERIVED
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 4400-0105P
; CURRENT APPLICATION NUMBER: US/10/145.012
; CURRENT FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-145-012-13

Query Match 39.5%; Score 325; DB 4; Length 614;
Best Local Similarity 99.8%; Pred. No. 2.7e-295;
Matches 445; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 378 YGVGRYQEVNPAITYITITFPFLFAMVFGDVGHGLMFLPALAMVLAENRPAVKAQNEIW 437
Db 162 YGVGRYQEVNPAITYITITFPFLFAMVFGDVGHGLMFLPALAMVLAENRPAVKAQNEIW 221
QY 438 QTFEFGRYLLMLGLFSIYTGFIYNECFSEATSIFFSGWSVAAMANQSGWSDAFLAQHTM 497
Db 222 QTFEFGRYLLMLGLFSIYTGFIYNECFSEATSIFFSGWSVAAMANQSGWSDAFLAQHTM 281
QY 498 LTLDPNVTGVFLGPYFPGIDPIWLSLAANHLSPNSFKMNSVILGVVHMAFGVVLGVFNH 557
Db 282 LTLDPNVTGVFLGPYFPGIDPIWLSLAANHLSPNSFKMNSVILGVVHMAFGVVLGVFNH 341
QY 558 VHFQGRHRLLETLPELTFLGLFGYLVFLVIYKMLCVWAARAAS-PSILIHFINMFLFS 616
Db 342 VHFQGRHRLLETLPELTFLGLFGYLVFLVIYKMLCVWAARAASAPSIILIHFINMFLFS 401
QY 617 HSPSNRLLYPROEVQATLVVLALAMVPILLGLTPTLHLLHRRRLRRPADQENKAG 676
Db 402 HSPSNRLLYPROEVQATLVVLALAMVPILLGLTPTLHLLHRRRLRRPADQENKAG 461
QY 677 LLDPDASVNGWSDEKAGGLDDEEAEELVPSEVLVLMHQAIIHTIEFCLGCVSNSTASYLRL 736
Db 462 LLDPDASVNGWSDEKAGGLDDEEAEELVPSEVLVLMHQAIIHTIEFCLGCVSNSTASYLRL 521
QY 737 WALSLAHQAQLSEVLWAMVMRIGLGLGREGVAAVLPVPIFAAFVAVMTVAILLVMEGLSAP 796
Db 522 WALSLAHQAQLSEVLWAMVMRIGLGLGREGVAAVLPVPIFAAFVAVMTVAILLVMEGLSAP 581
QY 797 LHALRLHWHVEFQNKFSYGTGYKLSPP 822
Db 582 LHALRLHWHVEFQNKFSYGTGYKLSPP 607

```
RESULT 6
US-10-489-725-5
; Sequence 5, Application US/10489725
; Publication No. US20050048067A1
; GENERAL INFORMATION:
; APPLICANT: Utku, Nalan
; TITLE OF INVENTION: Peptides capable of modulating immune response
; FILE REFERENCE: GE19A46/P-WO
; CURRENT APPLICATION NUMBER: US/10/489,725
; CURRENT FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: 60/322,896
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/322,895
; PRIOR FILING DATE: 2001-09-17
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-489-725-5

Query Match      39.5%; Score 325; DB 5; Length 614;
Best Local Similarity 99.8%; Pred. No. 2.7e-295;
Matches 445; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 378 YGVGRYQEVNPAPYTIITPPFLFAVMFGDVGHGLLMFLFALAMVLAENRPAVKAQNEIW 437
DB 162 YGVGRYQEVNPAPYTIITPPFLFAVMFGDVGHGLLMFLFALAMVLAENRPAVKAQNEIW 221
QY 438 QTFFRGRYLLLMGLFSIYTGFIYNECFSRATSIIPSGWSVAAMANQSGWDAFLAQHTM 497
DB 222 QTFFRGRYLLLMGLFSIYTGFIYNECFSRATSIIPSGWSVAAMANQSGWDAFLAQHTM 281
QY 498 LTLDPNVTGVFLGPPYFGIDPIWSLAANHLSFLNSFKMKMSVILGVVHMAFGVVLGVFNH 557
DB 282 LTLDPNVTGVFLGPPYFGIDPIWSLAANHLSFLNSFKMKMSVILGVVHMAFGVVLGVFNH 341
QY 558 VHFQGRHRLLTLPBELTFLGLFGYLVFLVIYKMLCVWAARAAS-PSILIHFINNPLFS 616
DB 342 VHFQGRHRLLTLPBELTFLGLFGYLVFLVIYKMLCVWAARAASAPSILIHFINNPLFS 401
QY 617 HSPSNRLLYPROEVVQATLVVLALAMVPILLGLTPLHLHRRHRRRLRRPADRQENKAG 676
DB 402 HSPSNRLLYPROEVVQATLVVLALAMVPILLGLTPLHLHRRHRRRLRRPADRQENKAG 461
QY 677 LLDLPDASVNGWSSDEEKAGGLDDEEAELVPSEVLMHQAIHTIEFCLGCVSNTASYLRL 736
DB 462 LLDLPDASVNGWSSDEEKAGGLDDEEAELVPSEVLMHQAIHTIEFCLGCVSNTASYLRL 521
QY 737 WALSLAHQAQLSEVLWAMVNRIGLIGRGVGAAVLVPIFAAFVMTVAILLVMEGLSAP 796
DB 522 WALSLAHQAQLSEVLWAMVNRIGLIGRGVGAAVLVPIFAAFVMTVAILLVMEGLSAP 581
QY 797 LHALRLHWVEFQNKFSYGTGYKLSPP 822
DB 582 LHALRLHWVEFQNKFSYGTGYKLSPP 607

RESULT 7
US-11-126-866-13
; Sequence 13, Application US/11126866
; Publication No. US20050220789A1
; GENERAL INFORMATION:
; APPLICANT: UTKU, et al., NALAN
; TITLE OF INVENTION: METHODS OF TREATMENT USING AN ANTIBODY THAT SPECIFICALLY BINDS TO
; FILE REFERENCE: 1472/71099-2A/JPW/AG
; CURRENT APPLICATION NUMBER: US/11/126,866
; CURRENT FILING DATE: 2005-05-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-126-866-13

Query Match      39.5%; Score 325; DB 6; Length 614;
Best Local Similarity 99.8%; Pred. No. 2.7e-295;
Matches 445; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 378 YGVGRYQEVNPAPYTIITPPFLFAVMFGDVGHGLLMFLFALAMVLAENRPAVKAQNEIW 437
```

```
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-126-866-13

Query Match      39.5%; Score 325; DB 6; Length 614;
Best Local Similarity 99.8%; Pred. No. 2.7e-295;
Matches 445; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 378 YGVGRYQEVNPAPYTIITPPFLFAVMFGDVGHGLLMFLFALAMVLAENRPAVKAQNEIW 437
DB 162 YGVGRYQEVNPAPYTIITPPFLFAVMFGDVGHGLLMFLFALAMVLAENRPAVKAQNEIW 221
QY 438 QTFFRGRYLLLMGLFSIYTGFIYNECFSRATSIIPSGWSVAAMANQSGWDAFLAQHTM 497
DB 222 QTFFRGRYLLLMGLFSIYTGFIYNECFSRATSIIPSGWSVAAMANQSGWDAFLAQHTM 281
QY 498 LTLDPNVTGVFLGPPYFGIDPIWSLAANHLSFLNSFKMKMSVILGVVHMAFGVVLGVFNH 557
DB 282 LTLDPNVTGVFLGPPYFGIDPIWSLAANHLSFLNSFKMKMSVILGVVHMAFGVVLGVFNH 341
QY 558 VHFQGRHRLLTLPBELTFLGLFGYLVFLVIYKMLCVWAARAAS-PSILIHFINNPLFS 616
DB 342 VHFQGRHRLLTLPBELTFLGLFGYLVFLVIYKMLCVWAARAASAPSILIHFINNPLFS 401
QY 617 HSPSNRLLYPROEVVQATLVVLALAMVPILLGLTPLHLHRRHRRRLRRPADRQENKAG 676
DB 402 HSPSNRLLYPROEVVQATLVVLALAMVPILLGLTPLHLHRRHRRRLRRPADRQENKAG 461
QY 677 LLDLPDASVNGWSSDEEKAGGLDDEEAELVPSEVLMHQAIHTIEFCLGCVSNTASYLRL 736
DB 462 LLDLPDASVNGWSSDEEKAGGLDDEEAELVPSEVLMHQAIHTIEFCLGCVSNTASYLRL 521
QY 737 WALSLAHQAQLSEVLWAMVNRIGLIGRGVGAAVLVPIFAAFVMTVAILLVMEGLSAP 796
DB 522 WALSLAHQAQLSEVLWAMVNRIGLIGRGVGAAVLVPIFAAFVMTVAILLVMEGLSAP 581
QY 797 LHALRLHWVEFQNKFSYGTGYKLSPP 822
DB 582 LHALRLHWVEFQNKFSYGTGYKLSPP 607

RESULT 8
US-11-126-841A-13
; Sequence 13, Application US/11126841A
; Publication No. US20050271659A1
; GENERAL INFORMATION:
; APPLICANT: UTKU, et al., NALAN
; TITLE OF INVENTION: METHODS OF TREATMENT USING AN ANTIBODY THAT SPECIFICALLY BINDS TO
; FILE REFERENCE: 1472/71099-2B/JPW/AG
; CURRENT APPLICATION NUMBER: US/11/126,841A
; CURRENT FILING DATE: 2005-05-10
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-11-126-841A-13

Query Match      39.5%; Score 325; DB 6; Length 614;
Best Local Similarity 99.8%; Pred. No. 2.7e-295;
Matches 445; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 378 YGVGRYQEVNPAPYTIITPPFLFAVMFGDVGHGLLMFLFALAMVLAENRPAVKAQNEIW 437
```

Db 162 YGGRYQEVNPAVYTIITFPFLFAVMFGDVCHGLLMFLFALAMVLAENRPAVKAQNEIW 221
QY 438 QTFPRGRYLLLMGLFSIYTGFIYNECFSRATSIPTPSGWSVAAMANOSGWSDAFLAQHTM 497
Db 222 QTFPRGRYLLLMGLFSIYTGFIYNECFSRATSIPTPSGWSVAAMANOSGWSDAFLAQHTM 281
QY 498 LTLDPNVTGVLGYPFGIDPIWLSAANHLFSFLNSFKMKMSVILGVVHMAFGVVLGVFNH 557
Db 282 LTLDPNVTGVLGYPFGIDPIWLSAANHLFSFLNSFKMKMSVILGVVHMAFGVVLGVFNH 341
QY 558 VHFQORHLLLETLPELTFLGLGFLYVFLVYKWLVCWMAAAS-PSILHFINMFLFS 616
Db 342 VHFQORHLLLETLPELTFLGLGFLYVFLVYKWLVCWMAAASPSILHFINMFLFS 401
QY 617 HSPSNRLLYPRQEVVQATLVIALAMPILLGTHPLHLLHRRRLRRRPPADROENKAG 676
Db 402 HSPSNRLLYPRQEVVQATLVIALAMPILLGTHPLHLLHRRRLRRRPPADROENKAG 461
QY 677 LLDLPDASVNGSSDEEKAGGLDDEBAELVPSVLMHQATHTEFCCLGCVSNTASYLRL 736
Db 462 LLDLPDASVNGSSDEEKAGGLDDEBAELVPSVLMHQATHTEFCCLGCVSNTASYLRL 521
QY 737 WALSLAQAQLSEVLWAMVMRIGLGLGREVGVAAYVLPVIFAFAVMTVAIILLVMEGLSAF 796
Db 522 WALSLAQAQLSEVLWAMVMRIGLGLGREVGVAAYVLPVIFAFAVMTVAIILLVMEGLSAF 581
QY 797 LHALRLHWVEFONKFYSGTGKLSPP 822
Db 582 LHALRLHWVEFONKFYSGTGKLSPP 607

RESULT 9

US-10-264-049-2656
; Sequence 2656, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Bitsee et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAI33PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2656
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (77)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (318)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (598)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-2656

Query Match 35.4%; Score 291; DB 4; Length 643;
Best Local Similarity 99.6%; Pred. No. 2.2e-263;
Matches 491; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 110 DVGNGQAQLRAQLHQLAHLVLRQGHPEPQLAAHTDGASERTPLLOAPGPGHQLRVNF 169
Db 78 DVGNGQAQLRAQLHQLAHLVLRQGHPEPQLAAHTDGASERTPLLOAPGPGHQLRVNF 137

QY 170 VAGAVEPHKAPALBRLLWRACRGFLIASPRELEQPLEHPVTGEBPATWMTFLISWGEQIG 229
Db 138 VAGAVEPHKAPALBRLLWRACRGFLIASPRELEQPLEHPVTGEBPATWMTFLISWGEQIG 197
QY 230 QKIRKTIQDCFHCHVFPFLQOEAEARLGAQLQOQSQOEVLQETERFLSOLGRVLQLL 289
Db 198 QKIRKTIQDCFHCHVFPFLQOEAEARLGAQLQOQSQOEVLQETERFLSOLGRVLQLL 257
QY 290 PPGQVQVHKMAVYLAALNQCSTVTHKCLIAEAWCSVRDLPALQEARLDRSSMEGVSAVA 349
Db 258 PPGQVQVHKMAVYLAALNQCSTVTHKCLIAEAWCSVRDLPALQEARLDRSSMEGVSAVA 317
QY 350 HRIPCRMPPTLRTNRFTASFOGIVDRYGVGRVQEVNPAPTYITITPPFLFAVMFGDVGH 409
Db 318 XRIPCRMPPTLRTNRFTASFOGIVDAYGVGRVQEVNPAPTYITITPPFLFAVMFGDVGH 377
QY 410 GLIMFLFALAMVLAENRPAVKAQNEIWQTFRRGRYLLLMGLFSIYTGFIYNECFSRAT 469
Db 378 GLIMFLFALAMVLAENRPAVKAQNEIWQTFRRGRYLLLMGLFSIYTGFIYNECFSRAT 437
QY 470 SIFPSGWSVAAMANOSGWSDAFLAQTMLTLDPNVTGVFLGYPFGIDPIWLSAANHLSP 529
Db 438 SIFPSGWSVAAMANOSGWSDAFLAQTMLTLDPNVTGVFLGYPFGIDPIWLSAANHLSP 497
QY 530 LNSFKMKMSVILGVVHMAFGVVLGVFNHVFHFGQRHLLLETLPELTFLGLGFLYVFLVI 589
Db 498 LNSFKMKMSVILGVVHMAFGVVLGVFNHVFHFGQRHLLLETLPELTFLGLGFLYVFLVI 557
QY 590 YKWLVCWMAAAS 602
Db 558 YKWLVCWMAAAS 570

RESULT 10

US-10-783-519-9
; Sequence 9, Application US/10783519
; Publication No. US20050064448A1
; GENERAL INFORMATION:
; APPLICANT: Stashenko, Philip
; APPLICANT: Li, Yi-Ping
; TITLE OF INVENTION: Osteoclast Proton Pump Subunit
; FILE REFERENCE: 1564.1006-001
; CURRENT APPLICATION NUMBER: US/10/783,519
; CURRENT FILING DATE: 2004-02-20
; PRIOR APPLICATION NUMBER: US/09/618,304
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 08/605,378
; PRIOR FILING DATE: 1996-02-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-783-519-9

Query Match 4.0%; Score 33; DB 5; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.2e-22;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 764 EVGVAAVVLVPIFAAFVMTVAIILLVMEGLSAF 796
Db 1 EVGVAAVVLVPIFAAFVMTVAIILLVMEGLSAF 33

RESULT 11

US-10-783-519-4
; Sequence 4, Application US/10783519
; Publication No. US2005006448A1
; GENERAL INFORMATION:
; APPLICANT: Stashenko, Philip
; APPLICANT: Li, Yi-Ping
; TITLE OF INVENTION: Osteoclast Proton Pump Subunit

FILE REFERENCE: 1564.1006-001
CURRENT APPLICATION NUMBER: US/10/783,519
CURRENT FILING DATE: 2004-02-20
PRIOR APPLICATION NUMBER: US/09/618,304
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 08/605,378
PRIOR FILING DATE: 1996-02-22
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 32
TYPE: PRT
ORGANISM: Homo sapiens
US-10-783-519-4

Query Match 3.9%; Score 32; DB 5; Length 32;
Best Local Similarity 100.0%; Pred. No. 9.9e-22;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 391 YTIITPPFLFVAVMGDVGHLLMFLFALAMVL 422
Db 1 YTIITPPFLFVAVMGDVGHLLMFLFALAMVL 32

RESULT 12
US-10-783-519-7
Sequence 7, Application US/10783519
Publication No. US2005006448A1
GENERAL INFORMATION:
APPLICANT: Stashenko, Philip
APPLICANT: Li, Yi-Ping
TITLE OF INVENTION: Osteoclast Proton Pump Subunit
FILE REFERENCE: 1564.1006-001
CURRENT APPLICATION NUMBER: US/10/783,519
CURRENT FILING DATE: 2004-02-20
PRIOR APPLICATION NUMBER: US/09/618,304
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 08/605,378
PRIOR FILING DATE: 1996-02-22
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 30
TYPE: PRT
ORGANISM: Homo sapiens
US-10-783-519-7

Query Match 3.6%; Score 30; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 7e-20;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 571 LPBLTFLGLGFLYFLVIYKWLVCWAARA 600
Db 1 LPBLTFLGLGFLYFLVIYKWLVCWAARA 30

RESULT 13
US-10-489-725-7
Sequence 7, Application US/10489725
Publication No. US20050048067A1
GENERAL INFORMATION:
APPLICANT: Utku, Nalan
TITLE OF INVENTION: Peptides capable of modulating immune response
FILE REFERENCE: GE19A46/P-WO
CURRENT APPLICATION NUMBER: US/10/489,725
CURRENT FILING DATE: 2004-03-12
PRIOR APPLICATION NUMBER: 60/322,896
PRIOR FILING DATE: 2001-09-17
PRIOR APPLICATION NUMBER: 60/322,895
PRIOR FILING DATE: 2001-09-17
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7

LENGTH: 29
TYPE: PRT
ORGANISM: Homo sapiens
US-10-489-725-7

Query Match 3.5%; Score 29; DB 5; Length 29;
Best Local Similarity 100.0%; Pred. No. 5.9e-19;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 664 RRPADQENKAGLLDLPDASVNGWSSDE 692
Db 1 RRPADQENKAGLLDLPDASVNGWSSDE 29

RESULT 14
US-10-783-519-5
Sequence 5, Application US/10783519
Publication No. US2005006448A1
GENERAL INFORMATION:
APPLICANT: Stashenko, Philip
APPLICANT: Li, Yi-Ping
TITLE OF INVENTION: Osteoclast Proton Pump Subunit
FILE REFERENCE: 1564.1006-001
CURRENT APPLICATION NUMBER: US/10/783,519
CURRENT FILING DATE: 2004-02-20
PRIOR APPLICATION NUMBER: US/09/618,304
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 08/605,378
PRIOR FILING DATE: 1996-02-22
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 26
TYPE: PRT
ORGANISM: Homo sapiens
US-10-783-519-5

Query Match 3.2%; Score 26; DB 5; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.5e-16;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 438 QTFEGRYLLLMGLFSIYTGFIYNE 463
Db 1 QTFEGRYLLLMGLFSIYTGFIYNE 26

RESULT 15
US-10-489-725-9
Sequence 9, Application US/10489725
Publication No. US20050048067A1
GENERAL INFORMATION:
APPLICANT: Utku, Nalan
TITLE OF INVENTION: Peptides capable of modulating immune response
FILE REFERENCE: GE19A46/P-WO
CURRENT APPLICATION NUMBER: US/10/489,725
CURRENT FILING DATE: 2004-03-12
PRIOR APPLICATION NUMBER: 60/322,896
PRIOR FILING DATE: 2001-09-17
PRIOR APPLICATION NUMBER: 60/322,895
PRIOR FILING DATE: 2001-09-17
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 24
TYPE: PRT
ORGANISM: Homo sapiens
US-10-489-725-9

Query Match 2.9%; Score 24; DB 5; Length 24;
Best Local Similarity 100.0%; Pred. No. 2.4e-14;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 679 DLPDASVNGWSSDEEKAGLDDEE 702

Db 1 DLPDASVNGWSDEERKAGGLDDEE 24

Search completed: June 29, 2006, 13:27:06
Job time : 189 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2006, 13:24:10 ; Search time 21 Seconds
(without alignments)
924.542 Million cell updates/sec

Title: US-10-783-519-2

Perfect score: 822

Sequence: 1 MGSMFRSEVALVQLFLPTA.....HWVEFQNKFGYGTGYKLSPF 822

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 103426 seqs, 23619683 residues

Word size : 1

Total number of hits satisfying chosen parameters: 103410

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA New.*

- 1: /EMC_Celerra_SID33/prodata/1/pubpaa/US09_NEW_PUB.pep.*
- 2: /EMC_Celerra_SID33/prodata/1/pubpaa/US06_NEW_PUB.pep.*
- 3: /EMC_Celerra_SID33/prodata/1/pubpaa/US07_NEW_PUB.pep.*
- 4: /EMC_Celerra_SID33/prodata/1/pubpaa/US08_NEW_PUB.pep.*
- 5: /EMC_Celerra_SID33/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 6: /EMC_Celerra_SID33/prodata/1/pubpaa/US10_NEW_PUB.pep.*
- 7: /EMC_Celerra_SID33/prodata/1/pubpaa/US11_NEW_PUB.pep.*
- 8: /EMC_Celerra_SID33/prodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|---------------------|-------------------|
| 1 | 545 | 66.3 | 830 | US-10-511-937-2474 | Sequence 2474, Ap |
| 2 | 20 | 2.4 | 43 | US-10-449-902-29221 | Sequence 29221, A |
| 3 | 20 | 2.4 | 271 | US-10-953-349-18081 | Sequence 18081, A |
| 4 | 20 | 2.4 | 273 | US-10-953-349-18080 | Sequence 18080, A |
| 5 | 20 | 2.4 | 314 | US-10-953-349-18079 | Sequence 18079, A |
| 6 | 20 | 2.4 | 820 | US-10-449-902-54048 | Sequence 54048, A |
| 7 | 17 | 2.1 | 357 | US-10-449-902-29928 | Sequence 29928, A |
| 8 | 17 | 2.1 | 489 | US-10-449-902-52773 | Sequence 52773, A |
| 9 | 17 | 2.1 | 818 | US-10-449-902-47207 | Sequence 47207, A |
| 10 | 17 | 2.1 | 818 | US-10-449-902-53444 | Sequence 53444, A |
| 11 | 17 | 2.1 | 923 | US-10-449-902-41603 | Sequence 41603, A |
| 12 | 8 | 1.0 | 98 | US-10-449-902-50960 | Sequence 50960, A |
| 13 | 8 | 1.0 | 136 | US-10-953-349-32586 | Sequence 32586, A |
| 14 | 8 | 1.0 | 301 | US-10-449-902-32058 | Sequence 32058, A |
| 15 | 8 | 1.0 | 335 | US-10-449-902-33155 | Sequence 33155, A |
| 16 | 8 | 1.0 | 444 | US-10-449-902-31385 | Sequence 31385, A |
| 17 | 8 | 1.0 | 518 | US-10-449-902-37736 | Sequence 37736, A |
| 18 | 8 | 1.0 | 623 | US-10-449-902-50362 | Sequence 50362, A |
| 19 | 8 | 1.0 | 678 | US-10-449-902-55991 | Sequence 55991, A |
| 20 | 8 | 1.0 | 777 | US-10-449-902-46537 | Sequence 46537, A |
| 21 | 8 | 1.0 | 777 | US-10-449-902-56066 | Sequence 56066, A |
| 22 | 7 | 0.9 | 106 | US-10-953-349-28780 | Sequence 28780, A |
| 23 | 7 | 0.9 | 117 | US-10-953-349-28801 | Sequence 28801, A |
| 24 | 7 | 0.9 | 120 | US-10-449-902-49220 | Sequence 49220, A |
| 25 | 7 | 0.9 | 149 | US-10-953-349-18110 | Sequence 18110, A |

| | | | | | | |
|----|---|-----|-----|---|---------------------|-------------------|
| 26 | 7 | 0.9 | 153 | 6 | US-10-953-349-35214 | Sequence 35214, A |
| 27 | 7 | 0.9 | 161 | 6 | US-10-953-349-27011 | Sequence 27011, A |
| 28 | 7 | 0.9 | 167 | 6 | US-10-449-902-43929 | Sequence 43929, A |
| 29 | 7 | 0.9 | 168 | 6 | US-10-449-902-49250 | Sequence 49250, A |
| 30 | 7 | 0.9 | 170 | 6 | US-10-505-928-861 | Sequence 861, App |
| 31 | 7 | 0.9 | 176 | 7 | US-11-293-697-3250 | Sequence 3250, Ap |
| 32 | 7 | 0.9 | 182 | 6 | US-10-449-902-54303 | Sequence 54303, A |
| 33 | 7 | 0.9 | 187 | 6 | US-10-449-902-49621 | Sequence 49621, A |
| 34 | 7 | 0.9 | 189 | 6 | US-10-953-349-32783 | Sequence 32783, A |
| 35 | 7 | 0.9 | 190 | 6 | US-10-953-349-25429 | Sequence 25429, A |
| 36 | 7 | 0.9 | 193 | 6 | US-10-449-902-49436 | Sequence 49436, A |
| 37 | 7 | 0.9 | 210 | 7 | US-11-075-398-7 | Sequence 7, Appli |
| 38 | 7 | 0.9 | 212 | 6 | US-10-449-902-39836 | Sequence 39836, A |
| 39 | 7 | 0.9 | 216 | 6 | US-10-953-349-25897 | Sequence 25897, A |
| 40 | 7 | 0.9 | 220 | 6 | US-10-449-902-45122 | Sequence 45122, A |
| 41 | 7 | 0.9 | 223 | 6 | US-10-449-902-50795 | Sequence 50795, A |
| 42 | 7 | 0.9 | 226 | 6 | US-10-953-349-31843 | Sequence 31843, A |
| 43 | 7 | 0.9 | 245 | 6 | US-10-449-902-51856 | Sequence 51856, A |
| 44 | 7 | 0.9 | 257 | 6 | US-10-953-349-31818 | Sequence 31818, A |
| 45 | 7 | 0.9 | 262 | 6 | US-10-449-902-31992 | Sequence 31992, A |

ALIGNMENTS

RESULT 1

US-10-511-937-2474
; Sequence 2474, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Lv, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, MacDonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; CURRENT FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2474
; LENGTH: 830
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2474

Query Match 66.3%; Score 545; DB 6; Length 830;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 765; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

| | | | | | |
|----|-----|------------------------|---------------------------------------|------------------------|-----|
| QY | 57 | RCELEKTTFLQEEVRRAGLVLP | PPKGRLPAPP | PRDLRIQETERLAQELRDVRNQ | 116 |
| DB | 57 | RCELEKTTFLQEEVRRAGLVLP | PPKGRLPAPP | PRDLRIQETERLAQELRDVRNQ | 116 |
| QY | 117 | ALRAQLHQLHQAALVROGHEPQ | LAHAHTDGAERTFLLOAPGPHODLRVNFAGAVEP | 176 | |
| DB | 117 | ALRAQLHQLHQAALVROGHEPQ | LAHAHTDGAERTFLLOAPGPHODLRVNFAGAVEP | 176 | |
| QY | 177 | HKAPALERLLWRACRGFLTAS | FRLEQPLEHPVTGEPATWMTFLISYWGQIQGKIRKIT | 236 | |
| DB | 177 | HKAPALERLLWRACRGFLTAS | FRLEQPLEHPVTGEPATWMTFLISYWGQIQGKIRKIT | 236 | |

QY 237 DCFHCHVFPFLOEQEALGALQOQSOBELQVLGETERFLSQVLGRVLQLLPPGQOV 296
Db 237 DCFHCHVFPFLOEQEALGALQOQSOBELQVLGETERFLSQVLGRVLQLLPPGQOV 296
QY 297 HKMKAVYLLNQCSTVTHKCLIAEAWCSVRDLPALQELAROSSMEEGVSAVAHRIPCRD 356
Db 297 HKMKAVYLLNQCSTVTHKCLIAEAWCSVRDLPALQELAROSSMEEGVSAVAHRIPCRD 356
QY 357 MPPTLLTRNRTASFOGIVDRGVGVQVQVNPAPYTIITFPFLFAVMFGDVGHLLMFLP 416
Db 357 MPPTLLTRNRTASFOGIVDAVGRIQVQVNPAPYTIITFPFLFAVMFGDVGHLLMFLP 416
QY 417 ALAMVLAENRAPVAKAQNBIWQTFFRGRYLLIMGLFSIYTGFIYNECFSRATSIPPSGW 476
Db 417 ALAMVLAENRAPVAKAQNBIWQTFFRGRYLLIMGLFSIYTGFIYNECFSRATSIPPSGW 476
QY 477 SVAAANQSGWSDAFIAQHTMTLTLDPNVTGVFLGPPGIDPTIWSLAANHLSFLNSFKMK 536
Db 477 SVAAANQSGWSDAFIAQHTMTLTLDPNVTGVFLGPPGIDPTIWSLAANHLSFLNSFKMK 536
QY 537 MSVILGVVHMGVVLGVFNHVFQGHRLLETLPELTLFLLGLGVLVFLVYKWLVM 596
Db 537 MSVILGVVHMGVVLGVFNHVFQGHRLLETLPELTLFLLGLGVLVFLVYKWLVM 596
QY 597 AARAAS-PSILIHFINMFLFSHPNRLLYPROVVQATLVVLAMVPILLTGTPHLL 655
Db 597 AARAASPSILIHFINMFLFSHPNRLLYPROVVQATLVVLAMVPILLTGTPHLL 656
QY 656 HRHRRRLRRPADRQENKAGLLDLPDASVNGWSSDEEKAGGLDDEERAEVLPSEVLMHQ 715
Db 657 HRHRRRLRRPADRQENKAGLLDLPDASVNGWSSDEEKAGGLDDEERAEVLPSEVLMHQ 716
QY 716 AHTTEFCGCVSNTASYLRWALSIAHAQLSEVLWAMVMRIGLGRGVGAANVVLVPI 775
Db 717 AHTTEFCGCVSNTASYLRWALSIAHAQLSEVLWAMVMRIGLGRGVGAANVVLVPI 776
QY 776 FAFAVMTVAILLVMBGLSAFLHALRLHWVEFQNKFSYGTGYKLSPF 822
Db 777 FAFAVMTVAILLVMBGLSAFLHALRLHWVEFQNKFSYGTGYKLSPF 823

RESULT 2

US-10-449-902-29221
; Sequence 29221, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Research.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 29221
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-29221

Query Match 2.4%; Score 20; DB 6; Length 43;
Best Local Similarity 100.0%; Pred. No. 5.2e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 793 LSAPFLHALRLHWVEFQNKFY 812
Db 4 LSAPFLHALRLHWVEFQNKFY 23

RESULT 3

US-10-953-349-18081
; Sequence 18081, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 18081
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-18081

Query Match 2.4%; Score 20; DB 6; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.6e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 793 LSAPFLHALRLHWVEFQNKFY 812
Db 233 LSAPFLHALRLHWVEFQNKFY 252

RESULT 4

US-10-953-349-18080
; Sequence 18080, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 18080
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-18080

Query Match 2.4%; Score 20; DB 6; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.6e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 793 LSAPFLHALRLHWVEFQNKFY 812
Db 235 LSAPFLHALRLHWVEFQNKFY 254

RESULT 5

US-10-953-349-18079
; Sequence 18079, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: ENCODED THERBY
; CURRENT APPLICATION NUMBER: US/10/953,349
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 18079
; LENGTH: 314

; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-18079

Query Match 2.4%; Score 20; DB 6; Length 314;
Best Local Similarity 100.0%; Pred. No. 2.9e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 793 LSAPFLHALRLHWVEFQNKFY 812
|||||
Db 276 LSAPFLHALRLHWVEFQNKFY 295
|||||

RESULT 6
US-10-449-902-54048
; Sequence 54048, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54048
; LENGTH: 820
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-54048

Query Match 2.4%; Score 20; DB 6; Length 820;
Best Local Similarity 100.0%; Pred. No. 6.8e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 793 LSAPFLHALRLHWVEFQNKFY 812
|||||
Db 781 LSAPFLHALRLHWVEFQNKFY 800
|||||

RESULT 7
US-10-449-902-29928
; Sequence 29928, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29928
; LENGTH: 357
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-29928

Query Match 2.1%; Score 17; DB 6; Length 357;

Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 727 VSNTASYLRLWALS LAH 743
|||||
Db 257 VSNTASYLRLWALS LAH 273
|||||

RESULT 8
US-10-449-902-52773
; Sequence 52773, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 52773
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-52773

Query Match 2.1%; Score 17; DB 6; Length 489;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 727 VSNTASYLRLWALS LAH 743
|||||
Db 389 VSNTASYLRLWALS LAH 405
|||||

RESULT 9
US-10-449-902-47207
; Sequence 47207, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47207
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-47207

Query Match 2.1%; Score 17; DB 6; Length 818;
Best Local Similarity 100.0%; Pred. No. 5.2e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 727 VSNTASYLRLWALS LAH 743
|||||

Db 718 VNTASYLRRLWALS LAH 734

RESULT 10

US-10-449-902-53444
; Sequence 53444, Application US/10449902
; Publication No. US20060123505A1

; GENERAL INFORMATION:

; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.

; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF

; FILE REFERENCE: MOA-A0205Y1-US

; CURRENT APPLICATION NUMBER: US/10/449,902

; CURRENT FILING DATE: 2003-05-29

; PRIOR APPLICATION NUMBER: JP 2002-203269

; PRIOR FILING DATE: 2002-05-30

; PRIOR APPLICATION NUMBER: JP 2002-383870

; PRIOR FILING DATE: 2002-12-11

; NUMBER OF SEQ ID NOS: 56791

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 53444

; LENGTH: 818

; TYPE: PRT

; ORGANISM: Oryza sativa

US-10-449-902-53444

Query Match

Best Local Similarity 2.1%; Score 17; DB 6; Length 818;

Mismatches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 727 VNTASYLRRLWALS LAH 743

|||||

Db 718 VNTASYLRRLWALS LAH 734

RESULT 11

US-10-449-902-41603
; Sequence 41603, Application US/10449902
; Publication No. US20060123505A1

; GENERAL INFORMATION:

; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.

; APPLICANT: The Institute of Physical and Chemical Research.

; APPLICANT: Foundation for Advancement of International Science.

; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF

; FILE REFERENCE: MOA-A0205Y1-US

; CURRENT APPLICATION NUMBER: US/10/449,902

; CURRENT FILING DATE: 2003-05-29

; PRIOR APPLICATION NUMBER: JP 2002-203269

; PRIOR FILING DATE: 2002-05-30

; PRIOR APPLICATION NUMBER: JP 2002-383870

; PRIOR FILING DATE: 2002-12-11

; NUMBER OF SEQ ID NOS: 56791

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 41603

; LENGTH: 923

; TYPE: PRT

; ORGANISM: Oryza sativa

US-10-449-902-41603

Query Match

Best Local Similarity 2.1%; Score 17; DB 6; Length 923;

Mismatches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 727 VNTASYLRRLWALS LAH 743

|||||

Db 824 VNTASYLRRLWALS LAH 840

RESULT 12

US-10-449-902-50960

; Sequence 50960, Application US/10449902
; Publication No. US20060123505A1

; GENERAL INFORMATION:

; APPLICANT: National Institute of Agrobiological Sciences.

; APPLICANT: Bio-oriented Technology Research Advancement Institution.

; APPLICANT: The Institute of Physical and Chemical Research.

; APPLICANT: Foundation for Advancement of International Science.

; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF

; FILE REFERENCE: MOA-A0205Y1-US

; CURRENT APPLICATION NUMBER: US/10/449,902

; CURRENT FILING DATE: 2003-05-29

; PRIOR APPLICATION NUMBER: JP 2002-203269

; PRIOR FILING DATE: 2002-05-30

; PRIOR APPLICATION NUMBER: JP 2002-383870

; PRIOR FILING DATE: 2002-12-11

; NUMBER OF SEQ ID NOS: 56791

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 50960

; LENGTH: 98

; TYPE: PRT

; ORGANISM: Oryza sativa

US-10-449-902-50960

Query Match

Best Local Similarity 1.0%; Score 8; DB 6; Length 98;

Mismatches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 659 RRLRRRP 666

|||||

Db 52 RRLRRRP 59

RESULT 13

US-10-953-349-32586
; Sequence 32586, Application US/10953349
; Publication No. US20060107345A1

; GENERAL INFORMATION:

; APPLICANT: ALEXANDROV, Nikolai et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

; FILE REFERENCE: ENCODED THERBY

; CURRENT APPLICATION NUMBER: US/10/953,349

; CURRENT FILING DATE: 2004-09-30

; NUMBER OF SEQ ID NOS: 40252

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 32586

; LENGTH: 136

; TYPE: PRT

; ORGANISM: Triticum aestivum

US-10-953-349-32586

Query Match

Best Local Similarity 1.0%; Score 8; DB 6; Length 136;

Mismatches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 597 AARASPS 604

|||||

Db 71 AARASPS 78

RESULT 14

US-10-449-902-32058
; Sequence 32058, Application US/10449902
; Publication No. US20060123505A1

; GENERAL INFORMATION:

; APPLICANT: National Institute of Agrobiological Sciences.

; APPLICANT: Bio-oriented Technology Research Advancement Institution.

; APPLICANT: The Institute of Physical and Chemical Research.

; APPLICANT: Foundation for Advancement of International Science.

; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF

; FILE REFERENCE: MOA-A0205Y1-US

; CURRENT APPLICATION NUMBER: US/10/449,902

; CURRENT FILING DATE: 2003-05-29

; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32058
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-32058

Query Match 1.0%; Score 8; DB 6; Length 301;
Best Local Similarity 100.0%; Pred. No. 9.9; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY 654 LLHRHRR 661
| | | | |
Db 13 LLHRHRR 20

RESULT 15

US-10-449-902-33155
; Sequence 33155, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33155
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-33155

Query Match 1.0%; Score 8; DB 6; Length 335;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 654 LLHRHRR 661
| | | | |
Db 13 LLHRHRR 20

Search completed: June 29, 2006, 13:27:33
Job time : 22 secs

THIS PAGE BLANK (USPTO)